Titre: Title:	H2 and H∞ Filtering for Nonlinear Singular Systems
Auteur: Author:	Mohammad Dikko S. Aliyu
Date:	2011
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Aliyu, M. D. S. (2011). H2 and H∞ Filtering for Nonlinear Singular Systems [Thèse de doctorat, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/566/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/566/
Directeurs de recherche: Advisors:	Michel Perrier, & Luc Baron
Programme: Program:	Génie mécanique

UNIVERSITÉ DE MONTRÉAL

$\mathcal{H}_2$ AND $\mathcal{H}_\infty$ FILTERING FOR NONLINEAR SINGULAR SYSTEMS

MOHAMMAD DIKKO S. ALIYU
DÉPARTEMENT DE GÉNIE MÉCANIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR (Ph.D.)
(GÉNIE MÉCANIQUE)
AVRIL 2011

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

$\mathcal{H}_2$ AND $\mathcal{H}_\infty$ FILTERING FOR NONLINEAR SINGULAR SYSTEMS

présentée par: ALIYU Mohammad Dikko S.

en vue de l'obtention du diplôme de: Philosophiæ Doctor

a été dûment acceptée par le jury d'examen constitué de:

M. MAYER René, Ph.D., président

M. PERRIER Michel, Ph.D., membre et directeur de recherche

M. BARON Luc, Ph.D., membre et codirecteur de recherche

M. SRINIVASAN Bala, Ph.D., membre

M. AGHDAM Amir, Ph.D., membre

Dedicated to my late father

Alh. Sulaiman Abubakar

ACKNOWLEDGMENTS

I wish to express my profound gratitude and indebtedness to Prof. M. Perrier for accepting to be my Research Director after the loss of Prof. Boukas (RiP) my former director whom we dearly miss. I am grateful to Prof. Mayer for streamlining my research program and helping me overcome the turmoil of the loss of Prof. Boukas. My appreciation also goes to Profs. L. Baron, B. Srinivasan and A. Aghdam for accepting to participate in my dissertation committee.

I wish to thank Mr. Ahmad Haidar and Mr. Mathieu Lozeau for the wonderful companionship we had during our period of stay at the École Polytechnique. In the same breath, I also wish to express my thanks to École Polytechnique de Montréal for giving me the chance to study and the superb facilities that it provided me for my research activities. The co-operation and support provided by the staff of the Mechanical Engineering Department, in particular, Mrs. Josée Dugas, Mrs. Carol Fraser, Mrs. Tanya Alleyne, Mr. Omar Toukal, and Mrs Martine Bénard is also hereby acknowledged.

Lastly, I wish to express my love and gratitude to my family for their patience and prayers during the preparation of this dissertation.

RÉSUMÉ

Dans les dernières années, les systèmes singuliers des équations différentielles ont carrément explosé puisqu'on les trouve dans plusieurs champs d'applications allant des systèmes électromécaniques en passant par des circuits électroniques, réacteurs chimiques et/ou biologiques ainsi que les systèmes d'écoulement des fluides. Dans cette thèse, deux classes des systèmes singuliers non linéaires seront considérer, en l'occurrence : (i) systèmes singuliers perturbés, (ii) systèmes généralisés ou systèmes algébro-différentielles. Les techniques $\mathcal{H}_2$ et $\mathcal{H}_\infty$ pour l'estimation de l'état de ces classes seront développés ainsi que des conditions suffisantes pour la résolution des problèmes en termes des équations d'Hamilton-Jacobi seront présentés. Deux systèmes, temps-continu et discrets, seront considérés et, pour plus de viabilité des résultats, des exemples pratiques seront présentés et résolus.

ABSTRACT

Singular systems of differential equations arise in many areas of science and technology, including electro-mechanical systems, electronic circuits, chemical and biological reactors, and fluid flow systems. In this thesis, two classes of singular nonlinear systems are considered; namely, (i) singularly perturbed systems, and (ii) generalized systems, or descriptor, or differential-algebraic systems. $\mathcal{H}_2$ and $\mathcal{H}_\infty$ techniques for state estimation of these classes of systems are developed, and sufficient conditions for the solvability of the problems in terms of Hamilton-Jacobi equations are presented. Both continuous-time and discrete-time systems are considered, and examples are presented to show the usefulness of the results.

TABLE OF CONTENTS

DEDICATION iii

ACKNOWLEDGMENTS iv

RÉSUMÉ v

ABSTRACT vi

TABLE OF CONTENTS vii

LIST OF FIGURES xi

CHAPTER 1 INTRODUCTION 1

1.1 Review of Classical Estimation Theory 3

1.1.1 Least-Squares Estimation 3

1.1.2 Minimum-Variance Estimation 5

1.1.3 Maximum Likelihood Estimation (MLE) 9

1.1.4 Bayesian Estimation 11

1.2 Linear Filtering for Dynamic Systems 12

1.2.1 Wiener-Kolmogorov Theory 12

1.2.2 Minimum-Variance (Kalman Filtering) for Linear Dynamic Systems . 14

1.2.3 Maximum Likelihood Estimation for Linear Dynamic Systems 16

1.2.4 Bayesian Estimation for Linear Dynamic Systems 20

1.3 Nonlinear Filtering 23

1.3.1 Extended Kalman Filters (EKFs) and Unscented Kalman Filters (UKF) 24

1.3.2 Maximum Likelihood Recursive Nonlinear Filtering 28

1.3.3 Bayesian Nonlinear Filtering and Particle Filters (PFs) 29

1.4 Conclusion 34

CHAPTER 2 LITERATURE REVIEW 35

2.1 Review of Kalman ($\mathcal{H}_2$)-Filtering for Linear Singularly-Perturbed Systems . 37

2.2 Review of ($\mathcal{H}_\infty$)-Filtering for Linear Singularly-Perturbed Systems 42
2.3 Review of $\mathcal{H}_2$ Filtering for Linear Descriptor Systems 48
2.4 Review of $\mathcal{H}_\infty$ Filtering for Linear Descriptor Systems 54
2.5 Motivations and Research Objectives 61
2.6 Conclusion 64

CHAPTER 3 $\mathcal{H}_2$ FILTERING FOR SINGULARLY-PERTURBED NONLINEAR SYSTEMS 65
3.1 $\mathcal{H}_2$ Filtering for Continuous-Time Systems 65
3.1.1 Problem Definition and Preliminaries 66
3.1.2 Decomposition Filters 68
3.1.3 Aggregate Filters 82
3.1.4 Push-Pull Configuration 83
3.1.5 Examples 85
3.2 $\mathcal{H}_2$ Filtering for Discrete-time Systems 87
3.2.1 Problem Definition and Preliminaries 87
3.2.2 Discrete Decomposition Filters 89
3.2.3 Discrete-time Aggregate Filters 102
3.2.4 Discrete-time Push-Pull Configuration 104
3.3 Examples 106
3.4 Conclusion 110

CHAPTER 4 $\mathcal{H}_\infty$ FILTERING FOR SINGULARLY-PERTURBED NONLINEAR SYSTEMS 111
4.1 $\mathcal{H}_\infty$ Filtering for Continuous-time Systems 112
4.1.1 Problem Definition and Preliminaries 112
4.1.2 Solution to the $\mathcal{H}_\infty$ Filtering Problem Using Decomposition Filters . 114
4.1.3 Aggregate $\mathcal{H}_\infty$ Filters 128
4.1.4 Push-Pull Configuration 131
4.1.5 Examples 132
4.2 $\mathcal{H}_\infty$ Filtering for Discrete-time Systems 133

4.2.1 Problem Definition and Preliminaries 133

4.2.2 Solution to the Discrete-time $\mathcal{H}_\infty$ Filtering Problem Using Decomposition Filters 135

4.2.3 Discrete-time Aggregate $\mathcal{H}_\infty$ Filters 150

4.2.4 Examples 154

4.3 Conclusion 154

CHAPTER 5 $\mathcal{H}_2$ FILTERING FOR NONLINEAR DESCRIPTOR SYSTEMS . . 156

5.1 $\mathcal{H}_2$ Filtering for Continuous-time Nonlinear Descriptor Systems 156

5.1.1 Problem Definition and Preliminaries 157

5.1.2 $\mathcal{H}_2$ Singular Filters 159

5.1.3 $\mathcal{H}_2$ Normal Filters 164

5.1.4 The General case 170

5.1.5 Examples 173

5.2 $\mathcal{H}_2$ Filtering for Discrete-time Descriptor Systems 175

5.2.1 Problem Definition and Preliminaries 175

5.2.2 Solution to the $\mathcal{H}_2$ Filtering Problem Using Singular Filters 178

5.2.3 Discrete-time $\mathcal{H}_2$ Normal Filters 184

5.2.4 The General Discrete-time case 189

5.2.5 Examples 192

5.3 Conclusion 192

CHAPTER 6 $\mathcal{H}_\infty$ FILTERING FOR DESCRIPTOR NONLINEAR SYSTEMS . . 194

6.1 $\mathcal{H}_\infty$ Filtering for Continuous-time Systems 194

6.1.1 Problem Definition and Preliminaries 194

6.1.2 $\mathcal{H}_\infty$ Singular Filters 197

6.1.3 $\mathcal{H}_\infty$ Normal Filters 203

6.2 Examples 209

6.3 $\mathcal{H}_\infty$ Filtering for Discrete-time Systems 212

6.3.1 Problem Definition and Preliminaries 212

6.3.2 Discrete-time $\mathcal{H}_\infty$ Singular Filters 216

6.3.3 Discrete-time $\mathcal{H}_\infty$ Normal Filters 222
6.4 Examples 228
6.5 Conclusion 229

SUMMARY AND CONCLUSION 230

REFERENCES 232

LIST OF FIGURES

Figure 3.1 Set-up for $\mathcal{H}_2$ Filtering 67

Figure 3.2 Set-up for discrete-time $\mathcal{H}_2$ filtering 88

Figure 3.3 Tunnel-diode circuit. 107

Figure 3.4 Actual state and state estimate for Reduced aggregate $\mathcal{H}_2$ filter with unknown initial condition and noise variance 0.1. 108

Figure 3.5 Estimation error for Reduced aggregate $\mathcal{H}_2$ filter with unknown initial condition and noise variance 0.1. 108

Figure 4.1 Set-up for $\mathcal{H}_\infty$ Filtering 113

Figure 4.2 Set-up for discrete-time $\mathcal{H}_\infty$ filtering 134

Figure 5.1 Set-up for $\mathcal{H}_2$ Filtering 158

Figure 5.2 $\mathcal{H}_2$ singular filter performance with unknown initial condition and noise variance 0.2 174

Figure 5.3 $\mathcal{H}_2$ singular filter performance with unknown initial condition and noise variance 0.5 174

Figure 5.4 Set-up for $\mathcal{H}_2$ Filtering 177

Figure 6.1 Set-up for $\mathcal{H}_\infty$ Filtering 196

Figure 6.2 $\mathcal{H}_\infty$ singular filter performance with unknown initial condition and uniformly distributed noise with variance 0.2 211

Figure 6.3 $\mathcal{H}_\infty$ singular filter performance with unknown initial condition and $\mathcal{L}_2$-bounded disturbance 211

Figure 6.4 Set-up for Discrete-time $\mathcal{H}_\infty$ Filtering 213

CHAPTER 1

INTRODUCTION

The problem of determining the internal behavior or "state" of a system from noisy measurements is known as *"state estimation"* or *"filtering"*. Linear estimation dates back to Gauss (1795) and the development of the method of "least-squares". A very good account of the subject can be found in the following references (Anderson, 1979), (Grewal, 1993), (Sorenson, 1985). However, since the turn of the century, the work of Nobert Wiener and of Kalman R. E. have dominated the subject. They pioneered the application of statistical ideas which started with the work of Wiener and Kolmogorov (Wiener, 1949), (Kolmogorov, 1949) to filtering problems. A review of these two approaches are given at a later section. Nonlinear filtering theory however, is a more recent and evolving subject, and is still a challenging research area because it is richer and more involved than linear filtering.

Following the fundamental work of Kalman and Bucy (Kalman, 1960), (Kalman, 1961) in linear filtering theory in 1960-1961, a host of publications appeared, formally deriving various approaches to linear filtering algorithms using "least-squares" or "minimum mean-squares", "maximum-likelihood", and other Bayesian and classical statistical methods. These statistical methods were also formally applied to the nonlinear estimation problem using linearization of one sort or another and Kalman-like algorithms. This also led to the development of the extended Kalman-filter (EKF). These works received great financial support and impetus from the aerospace industries as well as the Navy and Air-force offices of scientific research in the USA. These industries and research organizations also spear-headed the application of these techniques to submarine and aircraft navigation, space flight (including the Ranger, Mariner and Apollo missions), as well as satellite orbit determination and navigation.

Furthermore, while Kalman and Bucy were formulating the statistical linear filtering theory in the United States, Stratonovich (Stratonovich, 1960) was developing the probabilistic approach to discrete-time nonlinear filtering theory in Russia. Later Kushner (Kushner, 1967a), (Kushner, 1964a) and Wonham (Wonham, 1963) indepedently developed the contnuous-time theory, and subsequently, Ho and Lee (Ho, 1964) and Jazwinski (Jazwinski, 1998) applied the probabilistic theory to discrete-time problems. Thereafter, most of the developments in the nonlinear theory were made by Kushner (Kushner, 1967b)-(Kushner, 1970).

In this Dissertation, we shall focus on the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ filtering techniques for nonlinear singular systems. Even though $\mathcal{H}_2$ and $\mathcal{H}_\infty$ filtering techniques have been applied to linear singular systems by many authors (see Chapter 2 for a review), to the best of our knowledge, the nonlinear problem has not received any attention. Therefore, we propose to discuss these problems in this Dissertation. We shall present new results for $\mathcal{H}_2$ and $\mathcal{H}_\infty$ filtering for nonlinear singular systems, and also specialize these results to the linear case.

Singular systems are classified into two main classes; namely, (i) singularly perturbed systems, and (ii) differential-algebraic systems, descriptor or generalized state-space systems. They are characterized by a singular parameter or matrix on the left hand side of the system differential or difference equation. Therefore their analysis and control becomes more complicated than regular systems.

The Dissertation is organized as follows. In the remainder of this chapter, we shall introduce notations and then give a review of the classical (deterministic, statistical and probabilistic) approaches to linear and nonlinear filtering theory as applied to dynamic systems. Then, in Chapter 2, we present a literature review of deterministic finite-dimensional, mainly $\mathcal{H}_2$ and $\mathcal{H}_\infty$ filtering theory for linear singular systems. This is followed by a presentation of our research objectives. Our research contribution starts in Chapter 3, where we shall present new results on $\mathcal{H}_2$-filtering for both continuous-time and discrete-time nonlinear singularly-perturbed systems. This is followed in Chapter 4 with a counterpart solution to the $\mathcal{H}_\infty$ problem for the same class of systems, and in both continuous-time and discrete-time. Then in Chapters 5 and 6, we present similar solutions to the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ filtering problems for nonlinear descriptor nonlinear systems respectively, and in both continuous-time and discrete-time. Finally, we give a brief conclusion in Chapter 7.

The notation is fairly standard except where otherwise stated. Moreover, $\|(.)\|$, will denote the Euclidean vector norm on $\Re^n$, while $\mathsf{E}\{.\}$ and $p(.)$ will denote respectively the mathematical expectation operator and probability measure. Other notations will be defined accordingly.

1.1 Review of Classical Estimation Theory

In this Section, we give a review of classical static estimation theory beginning with the least-squares method. Then in Section 1.2, we discuss the extensions of the above methods to linear dynamic systems including the Wiener-Kolmogorov theory and Kalman filtering theory. Finally, in Section 1.3, we review some of the statistical approaches to nonlinear filtering theory and the Stratonovich-Kushner theory.

1.1.1 Least-Squares Estimation

The earliest motivations for the development of estimation theory apparently originated from astronomical studies in which planet and comet motion was studied using telescopic measurements. The motion of these bodies can be completely characterized by six parameters, and to determine these parameters, telescopic measurements are taken. The problem then was to estimate the values of these parameters from these measurements. To solve this problem, a young revolutionary, Karl Friedrich Gauss, then 18 years old, developed the method of *least squares* in 1795. This method which was published in his book "*Theoria Motus Corporum Coelestium*" or the "Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections (Crassidis, 2004)" in 1809, is very simple and intuitive. However, the method was also independently discovered by Legendre in 1806 and he published his results in his book "*Nouvelles méthodes pour la determination des orbites des cométes.* The delay in Gauss's publication of his results is what led to the controversy of the original inventor. Gauss also predicted the maximum-likehood method which was later discovered by R. A. Fisher in 1912.

To review the least squares method, consider an ensemble $\mathcal{Y}$ of observations, or time measurements of a variable $y(t)$ given by

$$\mathcal{Y} = \{y(t_1), y(t_2), \ldots, y(t_m)\}$$

where $y(t)$ is a linear function of another variable $x(t)$. Denoting now the emsemble $\mathcal{Y}$ by a vector $y \in \Re^m$, and assuming that y is linearly related to x by the relation

$$y = Hx, \quad H \in \Re^{m \times n},$$

where x denotes the vector of values of $x(t)$. The problem then is: *find an estimate $\hat{x}$ of x from the ensemble $\mathcal{Y}$ such that, the sum of squares of the total errors*

$$J_1 = \frac{1}{2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^T(y_i - \hat{y}_i) \triangleq \frac{1}{2}(y - H\hat{x})^T(y - H\hat{x}) \triangleq \frac{1}{2}\|e\|^2 \tag{1.1}$$

is minimized, where $e = (e_1, \ldots, e_n)^T = y - H\hat{x}$ is the error vector. The solution to this problem is obtained by applying the necessary principle of optimality:

$$\nabla J_{1,\hat{x}} = -H^T(y - H\hat{x}) = 0 \Rightarrow \hat{x} = (H^T H)^{-1}H^T y. \tag{1.2}$$

The above basic algorithm can also be modified by including weights on the measurements, especially if they are made with unequal precision. By modifying the cost function (1.1) as

$$J_2 = \sum_{i=1}^{n}(y_i - H\hat{x}_i)^T W(y_i - H\hat{x}_i), \tag{1.3}$$

where $W \in \Re^{m \times m}$ is a weighting matrix, the result is the following modified algorithm

$$\hat{x} = (H^T W H)^{-1}H^T W y, \tag{1.4}$$

which is also called the *weighted least-squares* method. Similarly, other variants of the algorithm including *constrained least-squares, nonlinear least-squares, and the Levenberg-Marquard* method also exist (Crassidis, 2004).

1.1.2 Minimum-Variance Estimation

This method is an enhancement of the method of Least-squares by introducing probability concepts in it. Minimum-variance gives the "best way" (in a probabilistic sense) to find an optimal estimate. Consider as in the previous subsection the linear observation model

$$y(t) = Hx(t) + v(t) \tag{1.5}$$

for the variable $x(t) \in \Re^n$, where $v(t) \in \Re^n$ is the measurement error vector. We can concieve of an estimate for $x(t)$ defined by

$$\hat{x}(t) = My(t) + \nu(t) \tag{1.6}$$

where $M \in \Re^{n\times m}$, $\nu(t) \in \Re^n$ are suitable weighting parameters. The objective is then to minimize the variance of each of the components $x_i, i = 1, \ldots, n$ of $x(t)$, i.e.,

$$J_i = \frac{1}{2}\mathsf{E}\{(x_i - \hat{x}_i)^2\}, \quad i = 1, \ldots, n. \tag{1.7}$$

It follows that, if the measurement errors $v(t) = 0$, then $x = \hat{x}$ and from (1.5), (1.6), we have

$$\hat{x} = MHx + \nu.$$

This implies that M and ν should satisfy

$$MH = I, \quad \nu = 0 \tag{1.8}$$

and the desired estimator has the form

$$\hat{x} = My. \tag{1.9}$$

Let us now define the error covariance matrix for an unbiased estimator as

$$P = \mathsf{E}\{(x-\hat{x})(x-\hat{x})^T\}. \tag{1.10}$$

Then the objective $J = \sum_i J_i$ above, can be redefined as the augmented cost function

$$J = Tr\left[\mathsf{E}\{(x-\hat{x})(x-\hat{x})^T\}\right] + Tr\left[\Lambda(I - MH)\right], \tag{1.11}$$

where Λ is a Lagrange multiplier matrix. Now using parallel axis theorem (for unbiased estimate)

$$\mathsf{E}\{(\hat{x}-x)(\hat{x}-x)^T\} = \mathsf{E}\{\hat{x}\hat{x}^T\} - \mathsf{E}\{x\}\mathsf{E}\{x\}^T$$

and substituting (1.6) in (1.9), we get using $E\{v\} = 0$,

$$\mathsf{E}\{\hat{x}\} = \mathsf{E}\{My\} = \mathsf{E}\{MHx + Mv\} = MHx. \tag{1.12}$$

Similarly, using $\mathsf{E}\{vv^T\} = R$ and the assumption that x and v are uncorrelated, i.e., $\mathsf{E}\{xv^T\} = \mathsf{E}\{vx^T\} = 0$, we obtain

$$\mathsf{E}\{\hat{x}\hat{x}^T\} = MHxx^TH^TM^T + MRM^T, \tag{1.13}$$

and

$$J = \frac{1}{2}Tr[MRM^T] + Tr[\Lambda(I - MH)].$$

Then, using the matrix derivative-identities

$$\frac{\partial}{\partial X}Tr(AXB) = A^TB^T, \quad \frac{\partial}{\partial X}Tr(XAX^T) = X(A + A^T),$$

and applying the necessary conditions for optimality of J with respect to M and Λ, we get

$$\Lambda^T = (H^TR^{-1}H)^{-1} \tag{1.14}$$

$$M = \Lambda^TH^TR^{-1} = (H^TR^{-1}H)^{-1}H^TR^{-1}. \tag{1.15}$$

Finally, using (1.9) we obtain the optimal unbiased estimate

$$\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} y, \tag{1.16}$$

which is referred to as the *Gauss-Markov* Theorem.

Remark 1.1.1. *The minimum-variance estimator is an unbiased estimator, i.e.* $\mathsf{E}\{\hat{x}\} = x$. *This can be shown as follows.* $\hat{x} = My = MHx + Mv$. *Then, using the fact that* $MH = I$, $\mathsf{E}\{v\} = 0$, *and taking expectations, the result follows. If on the other hand,* $\hat{x}$ *is biased, then the difference* $\mathsf{E}\{\hat{x}\} - x$ *is the bias in* $\hat{x}$.

The above algorithm can be refined to obtain improved estimates if a priori estimate $\hat{x}_a \in \Re^n$ of the variable x and covariance matrix Q are available. As in the previous case, we assume a linear model of the form

$$y = Hx + v \tag{1.17}$$

where v is zero-mean with covariance

$$Cov\{v\} = \mathsf{E}\{vv^T\} = R,$$

and assume the true state x is related to the a priori estimate as

$$\hat{x}_a = x + w, \tag{1.18}$$

where w is also zero-mean random vector with covariance

$$Cov\{w\} = \mathsf{E}\{ww^T\} = Q.$$

Similarly, we also assume that the measurement errors v and the a priori errors w are uncorrelated so that $E\{wv^T\} = 0$. Moreover, the objective is to estimate x as a linear combination of the measurements y and the a priori estimate $\hat{x}_a$ as

$$\hat{x} = My + N\hat{x}_a + \nu \tag{1.19}$$

where $M \in \Re^{n\times m}$, $N \in \Re^{n\times n}$, and $\nu \in \Re^n$ are design parameters, and are selected such that the variances of the estimates $\hat{x}_i, i = 1, \ldots, n$ from their true values x

$$\tilde{J}_i = \frac{1}{2}\mathsf{E}\left\{(\hat{x}_i - x_i)^2\right\}, \quad i = 1, \ldots, n$$

are minimized.

Again, if $\hat{x} = x$, then we should have from (1.17)

$$y = Hx, \quad v = 0.$$

Moreover, if in addition the a priori estimates are also perfect, i.e. $\hat{x}_a = x$, then $w = 0$ and (1.19) yields

$$x = MHx + Nx + \nu = (MH + N)x + \nu,$$

which implies

$$MH + N = I, \quad \nu = 0.$$

Thus, the desired estimator (1.19) has the form

$$\hat{x} = My + N\hat{x}_a. \tag{1.20}$$

Similarly, we can define the following augmented cost function as

$$\tilde{J} = \frac{1}{2}Tr[\mathsf{E}\{(x - \hat{x})(x - \hat{x})^T\}] + Tr[\tilde{\Lambda}(I - MH - N)], \tag{1.21}$$

where again $\tilde{\Lambda}$ is a Lagrangian multiplier. Then, using (1.18),(1.17) in (1.20), we have

$$\hat{x} = (MH + N)x + Mv + Nw. \tag{1.22}$$

Further, if we assume as before that x and v, w are uncorrelated with each other, (1.21) becomes

$$\tilde{J} = \frac{1}{2}Tr[MRM^T + NQN^T] + Tr[\tilde{\Lambda}(I - MH - N)]. \tag{1.23}$$

Applying now the necessary conditions for optimality of M, N and $\tilde{\Lambda}$, we have

$$MR - \tilde{\Lambda}^T H^T = 0,$$
$$NQ - \tilde{\Lambda}^T = 0,$$
$$I - MH - N = 0.$$

Finally, solving the above three equations for $\tilde{\Lambda}$, M, N, we get

$$\tilde{\Lambda}^T = (H^T R^{-1} H + Q^{-1})^{-1} \tag{1.24}$$
$$M = (H^T R^{-1} H + Q^{-1})^{-1} H^T R^{-1} \tag{1.25}$$
$$N = (H^T R^{-1} H + Q^{-1})^{-1} Q^{-1}. \tag{1.26}$$

1.1.3 Maximum Likelihood Estimation (MLE)

This method was invented by R. A. Fisher, a geneticist, in 1912. It yields estimates for the unknown quantities which maximize the probability of obtaining the observed set of data. Without loss of generality, one may consider the following Gaussian density function as a likelihood function

$$f(y;x) = \left(\frac{1}{2\pi\sigma^2}\right)^{m/2} \exp\left[-\sum_{i=1}^{m}(y_i - \mu)^2/(2\sigma^2)\right] \tag{1.27}$$

where $y \in \Re^m$ represents the measurement data, while $x \in \Re^n$ represents the estimated variable. However, it is often conveneint to deal with the logarithm of the above likelihood function in the form

$$\ln[f(y;x)] = -\frac{m}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{m}(y_i - \mu)^2. \tag{1.28}$$

Then, given the measurement information y, the problem is to find an estimate $\hat{x}$ which maximizes $f(y;x)$. The *likelihood-loss function* is also a probability density function (pdf), or a joint-density function, given by

$$L(y;x) = \Pi_{i=1}^{m} f_i(y;x). \tag{1.29}$$

Thus, the goal of the method is to find $\hat{x}$ such that the probability of obtaining the observations y is maximized. Moreover, since $\ln[L(y;x)]$ is a monotone function of $L(y;x)$, finding the x to maximize $\ln[L(y;x)]$ is equivalent to maximizing $L(y;x)$. Therefore, the necessary and sufficient conditions for the optimal estimate are respectively

$$\left\{\frac{\partial}{\partial x}\ln[L(y;x)]\right\}\bigg|_{\hat{x}} = 0 \tag{1.30}$$

$$\frac{\partial^2}{\partial x \partial x^T}\ln[L(y;x)] < 0. \tag{1.31}$$

Equation (1.30) is usually referred to as the *likelihood equation.* The method is best illustrated with an example.

Example 1.1.1. *Consider the Gaussian density function (1.27) and the problem of estimating $x = (\mu, \sigma^2)$ from measuremets y that is related to x by the pdf $f(y;x)$. Then, a natural choice for $L(y;x)$ in this case is $L(y;x) = f(y;x)$, and therefore $\ln[L(y;x)]$ is given by (1.28). Applying now the maximum-likelihood condition (1.30) for μ and σ^2, we obtain*

$$\begin{aligned} \left\{\tfrac{\partial}{\partial \mu}\ln[L(y;x)]\right\}\Big|_{\hat{\mu}} &= \tfrac{1}{\sigma^2}\textstyle\sum_{i=1}^{m}(y_i - \hat{\mu})^2 = 0 \\ &\Longrightarrow \hat{\mu} = \tfrac{1}{m}\textstyle\sum_{i=1}^{m} y_i \end{aligned} \tag{1.32}$$

$$\begin{aligned} \left\{\tfrac{\partial}{\partial \sigma^2}\ln[L(y;x)]\right\}\Big|_{\hat{\sigma^2}} &= -\tfrac{m}{2\hat{\sigma}^2} + \tfrac{1}{2\hat{\sigma}^4}\textstyle\sum_{i=1}^{m}(y_i - \hat{\mu})^2 = 0 \\ &\Longrightarrow \hat{\sigma}^2 = \tfrac{1}{m}\textstyle\sum_{i=1}^{m}(y_i - \mu)^2. \end{aligned} \tag{1.33}$$

Maximum likelihood estimation has several advantages, including firstly, the invariance principle, i.e., if $\hat{x}$ is a maximum likelihood estimate of x, then for any function $\phi(x)$, the maximum likelihood estimate of ϕ, is $\phi(\hat{x})$. Secondly, the estimation errors in a maximum likelihood estimate can be shown to be asymptotically Gaussian, regardless of the density function used.

1.1.4 Bayesian Estimation

In Bayesian estimation, the parameters to be estimated in the vector x are assumed to be random variables with some a priori probability distribution. This a priori information is combined with the measurement information y using a conditional density function which is known as the *"a posteriori distribution"* and Baye's rule to estimate the parameters. The conditional density is then given by

$$f(x|y) = \frac{f(y|x)f(x)}{f(y)}, \tag{1.34}$$

and thus both $f(y|x)$ and $f(x)$ must be known in order to use the method. Moreover, since y is known, $f(y)$ is a normalization for $f(x|y)$ in the above equation (1.34), and

$$f(y) = \int_{-\infty}^{\infty} f(y|x)f(x)dx.$$

If the above integral exists, then the a posterior density function $f(x|y)$ is said to be proper; otherwise it is said to be improper, and in this case, $f(y)$ is set to $f(y) = 1$.

The estimate $\hat{x}$ which maximizes the conditional density (1.34) is known as the maximum a posteriori (MAP) estimator. Since $f(y)$ is known, the above problem can be represented in logarithmic terms as that of maximizing the objective function

$$J_{MAP} = \ln[f(y|\hat{x})] + \ln[f(\hat{x})] = \ln[L(y|\hat{x})] + \ln[f(\hat{x})] \tag{1.35}$$

where $L(y|\hat{x}) = f(y|\hat{x})$ is a likelihood function. Thus, MAP is closely related to MLE in the following respect:

(a) if the a priori distribution $f(\hat{x})$ is uniform, then MAP is equivalent to MLE;

(b) MAP estimation has the same asymptotic consistency and efficiency of MLE;

(c) the MAP estimator converges to the MLE for large samples; and

(d) the MAP estimator also satisfies the invariance principle.

1.2 Linear Filtering for Dynamic Systems

In this section, we review some classical estimation techniques for linear dynamic systems defined by state-space models. Many of the methods are extensions of the static methods discussed in the previous section. For example, the Kalman filter is the dynamic version of the least-squares method and is also a minimum-variance estimator. We also review some of the historical developments in linear filtering in the early 20th century. As already mentioned, R. A. Fisher in 1912-1920 introduced the method of maximum-likelihood estimation which provided incentives for subsequent developments that culminated with the discovery of the Kalman filter.

1.2.1 Wiener-Kolmogorov Theory

Thereafter, Kolmogorov in 1941 (Kolmogorov, 1949) and Wiener (Wiener, 1949) in 1942 independently developed a linear minimum mean-square estimation technique that received considerable attention and provided the foundation for the Kalman filter theory. Historically, Wiener was led to develop his linear theory from the desire to find a rational design for fire control systems. To discuss the approach, consider a vector stochastic process $y(t) \in \Re^m, t \in (t_0, t_1) \subset \Re$, observed over an interval (t_0, t_1), and it is desired to estimate from these observations another process $x(t) \in \Re^n, t \in (t_0, t_1)$, which is related to $y(t)$ linearly. Wiener's work (Wiener, 1949) assumed that $x(t)$ and $y(t)$ are jointly wide-sense stationary ergodic processes with $t_0 = -\infty$ and $n = m = 1$. The result of this investigation led to the specification of the minimum-variance unbiased estimate $\hat{x}$ of x by its weighting function $w(\tau)$ in the form of the convolution

$$\hat{x}(t) = \int_{-\infty}^{t} w(t-s)y(s)ds, \tag{1.36}$$

where $w(.)$ satisfies the Wiener-Hopf integral equation

$$\mathsf{E}\{x(t)y^T(\tau)\} = \mathsf{E}\{\hat{x}(t)y^T(\tau)\}. \tag{1.37}$$

Wiener used Fourier transform methods and spectral-factorization to solve this equation.

Further, at about the same time that Wiener was developing his continuous-time linear theory, Kolmogorov was developing an analogous discrete-time theory (Kolmogorov, 1949). To review the approach, consider the problem of estimating a signal $x_k, k \in \mathbf{Z}$, which is possibly time-varying, from measurement data $(y_0, y_1, \ldots, y_n)$ where x_k and $y_i, i \in \mathbf{Z}$ are linearly related by some cross-cerrelation function. Denote the estimate of x_k using measurements up to y_k by $\hat{x}_{k/k}$. Then, Kolmogorov used the discrete convolution

$$\hat{x}_{k/k} = \sum_{i=0}^{k} H_{k,i} y_i, \tag{1.38}$$

where $H_{k,i}$ are the filter gains (or coefficients) which are to be chosen so that the mean-square errors are minimized, i.e., $H_{k,i}$ are chosen such that

$$J_k = \mathsf{E}[(x_k - \hat{x}_{k/k})^T (x_k - \hat{x}_{k/k})] \tag{1.39}$$

is minimized for $k = 0, \ldots,$. A necessary and sufficient condition for the existence of such minimizers is that the estimation error or innovation $e_{k/k} = x_k - \hat{x}_{k/k}$ is orthogonal to the measurement data, i.e.,

$$\mathsf{E}[e_{k/k} y_i^T] = 0, \quad i = 0, 1, \ldots, k, \tag{1.40}$$

holds. The above is the discrete Wiener-Hopf equation which is usually written as

$$\mathsf{E}[x_k y_i^T] = \sum_{j=0}^{k} H_{k,j} \mathsf{E}[y_j y_i^T], \quad i = 0, 1, \ldots, k. \tag{1.41}$$

This equation must be solved for the filter coefficients $H_{k,j}$, and can be represented in matrix form, whose solution should be straight-forward. However, the matrix inversion that is required becomes computationally impractical when k is large. To circumvent this, Wiener and Kolmogorov assumed $k_0 = -\infty$ instead of $k_0 = 0$, and the system to be stationary. The resulting equations can then be solved using spectral factorization. Unfortunately, the application of the Wiener-Kolmogorov theory was very limited, because the problem of synthesis remained practically unresolved. However, in the 1950's many investigators generalized the Wiener-Hopf equations to scalar nonstationary processes and finite observations intervals. They also introduced shaping filters in (1.41) in order to help the solvability of the equation. Similarly, since a new solution for the weighting coefficients for the filter must be generated for each k, J. W. Folin suggested a recursive approach for generating $x_{k/k}$ given a new measurement. Nevertheless, the main restrictions and drawbacks of the Wiener-Kolmogorov theory are the following:

(a) the processes must be stationary and ergodic;

(b) $t_0 = -\infty$;

(c) the spectral factorization solution of the Wiener-Hopf equation is not amenable to numerical computation even for rational spectra;

(d) the measurements or observations must be scalar processes, otherwise factorization of matrices must be considered; and finally,

(e) the physical realization of the processor determined by the filter coefficients, is far from trivial.

1.2.2 Minimum-Variance (Kalman Filtering) for Linear Dynamic Systems

Subsequently, in 1960, Kalman published his work on the discrete-time version of the Kalman filter (Kalman, 1960). But prior to this, Peter Swerling had published at the RAND Corporation a Memo in 1958 about a recursive procedure for orbit determination (Sorenson, 1985). Therefore, there was a squabble between Kalman and Swerling similar to the Gauss-Legendre squabble about who was first to discover the Kalman filter, with the former prevailing. Similarly, Stratonovich (Stratonovich, 1960) in the USSR also published at about the same time results that are equivalent to Kalman's work. But next, Kalman and Bucy together published the second paper (Kalman, 1961) on the continuous-time version of the theory and a complete solution to the linear filtering and prediction problem.

The Kalman-Bucy theory provided explicit synthesis of the minimum-variance unbiased estimate of the state of the signal $x(t)$ by showing that it satisfies a stochastic differential equation which is driven by the observations. The central idea of the theory was to replace the problem of solving the Wiener-Hopf equation with that of solving a matrix Riccati equation which is considerably simpler. Moreover, the solution of this matrix Riccati equation is the error covariance matrix of the optimal estimate, and with the advent of digital computers, the theory was applied successfully to countless problems in guidance, navigation, and orbit determination. These applications include also some of the most challenging space programs of that time, including the Mariner, the Ranger, Apollo, and the ill-fated Voyager. Other applications also included submarine detection, fire control, and practical schemes for detection (Sorenson, 1985).

At this point, we summarize the main results of the Kalman-Bucy solution to the Wiener filtering problem in discrete-time. Consider the linear state equations given by

$$\Sigma_{ldk} : \begin{cases} x_{k+1} &= A_{k+1,k}x_k + w_k, \quad x(0) = x_0 \\ y_k &= H_k x_k + v_k \end{cases} \tag{1.42}$$

where $x \in \Re^n$, $y \in \Re^m$, $\{w_k\}$, $\{v_k\}$ are independent white-noise processes with zero-mean and second-order statistics given by

$$\mathsf{E}\{v_i v_j^T\} = R_i\delta_{ij}, \quad \mathsf{E}\{w_i w_j^T\} = Q_i\delta_{ij}, \quad \mathsf{E}\{v_i w_j^T\} = 0 \quad \text{for all } \ i, j. \tag{1.43}$$

Similarly, the initial condition x_0 is also assumed to be a random vector with mean value $\hat{x}_{0|-1}$, covariance matrix $P_{0|-1}$ and uncorrelated with v_k and w_k respectively. An estimate $\hat{x}_{k/k}$ for x_k is to be determined from the measurements $\{y_k\}$ and possibly previous estimates, to minimize the mean-square error of the estimates, i.e,

$$J_k = \mathsf{E}[(x_k - \hat{x}_k)^T(x_k - \hat{x}_k)], \quad k = 1, \ldots \tag{1.44}$$

A sequential estimator which operates in a recursive manner combining new measurement information y_k and the best previous estimate $\hat{x}_{k-1/k-1}$ is also desired. The solution to this problem can be determined from the orthogonality principle (1.40). It is also very intuitive at this point to conjecture an estimator of the form

$$x_{k/k} = A_{k/k-1}\hat{x}_{k-1/k-1} + K_k[y_k - H_k A_{k/k-1}\hat{x}_{k-1/k-1}] \tag{1.45}$$

which is a linear combination of the predicted estimate in the absence of new data, and the residuals or innovation $r_k = y_k - H_k x_{k/k-1}$. The gain matrix K_k is chosen to minimize J_k and is given by

$$K_k = P_{k/k-1}H_k^T(H_k P_{k/k-1}H_k^T + R_k)^{-1}, \tag{1.46}$$

where the matrix $P_{k/k-1}$ is the covariance of the error in the predicted estimate and is given by

$$P_{k/k-1} = \mathsf{E}[(x_k - \hat{x}_{k/k-1})(x_k - \hat{x}_{k/k-1})^T] = A_{k/k-1}P_{k-1/k-1}A_{k/k-1}^T + Q_{k-1}, \tag{1.47}$$

while $P_{k/k}$ is the covariance of the error in the estimate $\hat{x}_{k/k}$, and is given by

$$P_{k/k} = \mathsf{E}[(x_k - \hat{x}_{k/k})(x_k - \hat{x}_{k/k})^T] = P_{k/k-1} - K_k H_k P_{k/k-1}. \tag{1.48}$$

These equations (1.45)-(1.48) represent the discrete-time Kalman filter equations and the solution to the filtering problem.

1.2.3 Maximum Likelihood Estimation for Linear Dynamic Systems

In this subsection, we review an approach to the maximum likelihood method (Rauch, 1965) for estimating the state of a linear dynamic system. We reconsider the model (1.42), (1.43) with R_i positive definite, and in addition, the initial condition x_0 is assumed to be a Gaussian distributed random vector with

$$\mathsf{E}\{x_0\} = \bar{x}_0, \quad \mathsf{E}\{x_0 x_0^T) = \bar{P}_0.$$

The problem is to find an estimate $\hat{x}_{k/N}, k = 0, \ldots, K$ of x_k from the observations $\{y_0, \ldots, y_N\}$ so as to minimize the following objective functional

$$J_{ML} = \sum_{k=0}^{K} l(x_0, \hat{x}_{0/N}; x_1, \hat{x}_{1/N} \ldots, x_K, \hat{x}_{K/N}), \tag{1.49}$$

for some loss function $l(.)$ of the variables. The problem is called (i) filtering, if $K = N$; (ii) prediction, if $K \geq N$; and (iii) smoothing if $K \leq N$. We shall present a solution to the filtering and prediction problems. In order to solve the problem, the distribution of interest is the joint distribution of $x_0, \ldots, x_K$ conditioned on $y_0, \ldots, y_N$ defined by

$$p(x_0, \ldots, x_K / y_0, \ldots, y_N).$$

If the loss function $l(.)$ is zero near $x_k = \hat{x}_{k/k}$ for $k = 0, \ldots, K$, and very large otherwise, then the optimum procedure is to use the joint maximum likelihood function or the logarithm of the above probability distribution. If on the other hand, the objective functional (1.49) above has the special form

$$J_{ML} = \sum_{k=0}^{K} l_k(x_k, \hat{x}_{k/N}), \tag{1.50}$$

then the distribution of interest is the marginal distribution of x_k conditioned on $\{y_0, \ldots, y_N\}$, i.e.,

$$p(x_k / y_0, \ldots, y_N),$$

which can be obtained from $p(x_0, \ldots, x_K / y_0, \ldots, y_N)$ by summing out $x_j, j \neq k$. Moreover, if $l_k(x_k, \hat{x}_{k/N})$ is zero near $x_k = \hat{x}_{k/N}$ and very large otherwise, then the optimum procedure to use is the marginal maximum likelihood function, or the logarithm of the above distribution. We shall employ this for determining the solution to the filtering and prediction problems.

Let $Y_k \triangleq \{y_0, \ldots, y_k\}$ and the estimate based on this data by $\hat{x}_{k/k}$. This is to be obtained by maximizing the density function $p(x_k/Y_k)$, which is equivalent to maximizing

$$L(x_k, Y_k) = \log\, p(x_k/Y_k) = \log\, p(x_k, Y_k) - \log\, p(Y_k). \tag{1.51}$$

Using the fact that v_k are independent, we have

$$p(x_k, Y_k) = p(y_k/x_k, Y_{k-1})p(x_k, Y_{k-1}) = p(y_k/x_k)p(x_k, Y_{k-1})p(Y_{k-1}). \tag{1.52}$$

Next, let $\hat{x}_{k-1/k-1}$, $\hat{x}_{k/k-1}$ be the estimates of x_{k-1}, x_k given Y_{k-1} respectively, and let $e_{k-1/k-1}$, $e_{k/k-1}$ be the corresponding estimation errors. In addition, define

$$Cov\{x_{k-1/k-1}\} = P_{k-1/k-1}, \quad Cov\{x_{k/k-1}\} = P_{k/k-1}.$$

Since $v_k, k = 1, \ldots, N$ are independent, then

$$\hat{x}_{k/k-1} = A_{k,k-1}\hat{x}_{k-1/k-1} \tag{1.53}$$

$$P_{k/k-1} = A_{k,k-1}P_{k-1/k-1}A_{k,k-1}^T + Q_{k-1} \tag{1.54}$$

give a solution of the prediction problem. Further, using (1.42), (1.43), it follows that the conditional random vector x_k given Y_{k-1} has mean and covariance

$$\mathsf{E}\{x_k/Y_{k-1}\} = \hat{x}_{k/k-1}, \quad Cov\{x_k/Y_{k-1}\} = P_{k/k-1}, \tag{1.55}$$

while the conditional vector y_k given x_k has

$$\mathsf{E}\{y_k/x_k\} = H_k x_k, \quad Cov\{y_k/x_k\} = R_k. \tag{1.56}$$

Substituting (1.55), (1.56) in (1.52) and using the fact that all the vectors are normally distributed, we have

$$\begin{aligned} p(x_k, Y_k) &= \frac{1}{\sqrt{(2\pi)^m |R_k|}} \exp\left\{-1/2\|y_k - H_k x_k\|^2_{R_k^{-1}}\right\} |P_{k/k-1}|^{-1/2} \times \\ &\quad \exp\left\{-1/2\|x_k - \hat{x}_{k/k-1}\|^2_{P^{-1}_{k/k-1}}\right\} p(Y_{k-1}). \end{aligned} \tag{1.57}$$

Further, substituting (1.57) in (1.52) and separating the terms in $L(.)$ that depend on x_k, and defining the marginal maximum likelihood estimation (MLE) objective function in terms of these terms, we get

$$J_{MLE} = \|y_k - H_k x_k\|^2_{R_k^{-1}} + \|x_k - \hat{x}_{k/k-1}\|^2_{P^{-1}_{k/k-1}}. \tag{1.58}$$

Finally, applying the necessary condition for the optimal estimate, $\left.\frac{\partial J_{MLE}}{\partial x_k}\right|_{\hat{x}_{k/k}} = 0$, yields

$$\hat{x}_{k/k} = (H_k^T R_k^{-1} H_k + P^{-1}_{k/k-1})(H_k^T R_k^{-1} y_k + P^{-1}_{k/k-1}\hat{x}_{k/k-1}), \tag{1.59}$$

which is the solution of the filtering problem. An alternative representation of this solution can be given by using the following matrix-inversion lemma.

Lemma 1.2.1. *Suppose $S^{-1}_{k+1} = S_k^{-1} + H_k^T R_k^{-1} H_k$ where S_k and R_k are symmetric and positive definite. Then S_{k+1} exists and is given*

$$S_{k+1} = S_k - S_k H_k^T (H_k S_k H_k^T + R_k)^{-1} H_k S_k. \quad \diamondsuit$$

Using the above lemma in (1.59), we have the following more computationally efficient representation

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + B_k(y_k - H_k \hat{x}_{k/k-1}) = A_{k,k-1}\hat{x}_{k-1/k-1} + B_k(y_k - H_k A_{k/k-1}\hat{x}_{k-1/k-1}) \tag{1.60}$$

where

$$B_k = P_{k/k-1} H_k^T (H_k P_{k/k-1} H_k^T + R_k)^{-1}.$$

Similarly, substituting (1.42) in (1.60) yields the error equation

$$e_{k/k} = (I - B_k H_k)[A_{k,k-1} e_{k-1/k-1} + w_{k-1}] - B_k v_k. \tag{1.61}$$

Moreover, since $e_{k-1/k-1}$, v_k and w_{k-1} are independent, it follows that

$$P_{k/k} = Cov\{e_{k/k}\} = (I - B_k H_k) P_{k/k-1}. \tag{1.62}$$

Equations (1.60)-(1.62) are the same as those derived by Kalman and presented in the previous section with $\hat{x}_{0/-1} = \bar{x}_0$ and $P_{0/-1} = \bar{P}_0$.

1.2.4 Bayesian Estimation for Linear Dynamic Systems

In this subsection we review an approach to Bayesian state estimation for linear dynamic systems (Ho, 1964). We consider the following linear time-invariant system model

$$\Sigma_{ld} : \begin{cases} x_{k+1} &= Ax_k + Bw_k, \quad x(0) = x_0 \\ y_k &= H_k x_k + v_k \end{cases} \tag{1.63}$$

where $x \in \Re^n$, $y \in \Re^m$, and w and v are independent white Gaussian random sequences with

$$\mathsf{E}\{v_k\} = \mathsf{E}\{w_k\} = 0, \quad Cov\{v_{k+1}\} = R, \quad Cov\{w_k\} = Q.$$

Let $Y_{k+1} = \{y_0, \ldots, y_{k+1}\}$ be a set of discrete measurements, and suppose

$$p(x_k/Y_k) \quad \text{is Gaussian,}$$
$$Cov\{x_k/Y_k\} = P_k,$$
$$p(w_k, v_{k+1}/x_k, Y_k) = p(w_k)p(v_{k+1}),$$

are known. The problem is to find the best estimate $\hat{x}_{k+1}$ of x_k from Y_{k+1} in some optimal sense which will be defined later. The Bayesian solution can be obtained in the following steps:

1. Evaluate $p(x_{k+1}/x_k)$; this can be done either experimentally or analytically from knowledge of $p(w_k, v_{k+1}/x_k)$, $p(x_k/Y_k)$ and (1.63).

2. Evaluate $p(y_{k+1}/x_k, x_{k+1})$; this is is also derived from $p(w_k, v_{k+1}/x_k)$ and (1.63).

3. Evaluate

$$p(x_{k+1}, y_{k+1}/Y_k) = \int p(y_{k+1}/Y_k, x_{k+1})p(x_{k+1}/x_k)p(x_k/Y_k)dx_k. \tag{1.64}$$

From this, the marginal density functions $p(x_{k+1}/Y_k)$ and $p(y_{k+1}/Y_k)$ can be directly evaluated.

4. Evaluate

$$\begin{aligned} p(x_{k+1}/Y_{k+1}) &= \frac{p(x_{k+1}, y_{k+1}/Y_k)}{p(y_{k+1}/Y_k)} \\ &= \frac{\int p(y_{k+1}/Y_k, x_{k+1})p(x_{k+1}/x_k)p(x_k/Y_k)dx_k}{\int\int p(y_{k+1}/Y_k, x_{k+1})p(x_{k+1}/x_k)p(x_k/Y_k)dx_{k+1}dx_k} \end{aligned} \tag{1.65}$$

from (1.64). Equation (1.65) is a functional-integral-difference equation governing the evolution of the a posteriori density function of the state.

5. Estimate of x_{k+1} can then be obtained from $p(x_{k+1}/Y_{k+1})$.

Applying the above steps 1-5 to the model (1.63), we have since w_k, v_{k+1} are not dependent on the state, (1.65) simplifies to

$$p(x_{k+1}/Y_{k+1}) = \frac{p(y_{k+1}/x_{k+1})}{p(x_{k+1}/Y_k)}p(x_{k+1}/Y_k). \tag{1.66}$$

By assumption, $p(x_{k+1}/Y_k)$ is Gaussian and independent of v_{k+1}. Hence,

$$\left.\begin{aligned} \mathsf{E}\{x_{k+1}/Y_k\} &= A\hat{x}_k \\ Cov\{x_{k+1}/Y_k\} &= AP_kA^T + BQB^T \triangleq P_{k+1}. \end{aligned}\right\} \tag{1.67}$$

Similarly, $p(y_{k+1}/Y_k)$ is Gaussian and

$$\left.\begin{aligned} \mathsf{E}\{y_{k+1}/Y_k\} &= HA\hat{x}_k \\ Cov\{y_{k+1}/Y_k\} &= HP_{k+1}H^T + R. \end{aligned}\right\} \tag{1.68}$$

Finally, $p(y_{k+1}/Y_k)$ is also Gaussian

$$\left.\begin{array}{rcl} p(y_{k+1}/x_{k+1}\} & = & Hx_{k+1} \\ Cov\{y_{k+1}/x_{k+1}\} & = & R. \end{array}\right\} \tag{1.69}$$

Combining (1.67)-(1.69) and using (1.66) one gets

$$\begin{aligned} p(x_{k+1}/Y_{k+1}) \;=\; & \frac{|HM_{k+1}H^T + R|^{1/2}}{(2\pi)^{n/2}|R|^{1/2}|M_{k+1}|^{1/2}} \exp\Big\{-1/2[(x_{k+1} - A\hat{x}_k)^T M_{k+1}^T (x_{k+1} - \\ & A\hat{x}_k) + (y_{k+1} - Hx_{k+1})^T R^{-1}(y_{k+1} - Hx_{k+1}) - (y_{k+1} - \\ & HA\hat{x}_{k+1})^T (HM_{k+1}H^T + R)^{-1}(y_{k+1} - HA\hat{x}_{k+1})]\Big\}. \end{aligned} \tag{1.70}$$

Completing the squares in the terms in $\{.\}$, we get

$$p(x_{k+1}/Y_{k+1}) = \frac{|HM_{k+1}H^T + R|^{1/2}}{(2\pi)^{n/2}|R|^{1/2}|M_{k+1}|^{1/2}} \exp\Big\{-1/2(x_{k+1} - \hat{x}_{k+1})^T P_{k+1}^{-1}(x_{k+1} - \hat{x}_{k+1})\Big\}, \tag{1.71}$$

where

$$\hat{x}_{k+1} = A\hat{x}_k + M_{k+1}H^T(HM_{k+1}H^T + R)^{-1}(y_{k+1} - HA\hat{x}_k), \tag{1.72}$$

$$P_{k+1}^{-1} = M_{k+1}^{-1} + H^T R^{-1} H, \tag{1.73}$$

or equivalently

$$P_{k+1} = M_{k+1} - M_{k+1}H^T(HM_{k+1}H^T + R)^{-1}HM_{k+1}, \tag{1.74}$$

and

$$M_{k+1} = AP_kA^T + BQB^T. \tag{1.75}$$

Equations (1.72)-(1.75) are exactly the same as the Kalman filter equations presented in the previous subsection.

1.3 Nonlinear Filtering

In this section, we review briefly the history of nonlinear filtering and present some well-known approaches. Most of the results we present in this section will be in continuous-time, by virtue of the nature of the original publications.

The theory of nonlinear filtering also started in the early 1960s, and was originally developed by Stratonovich. He formally obtained the random partial differential equation for the conditional density $p(x_t|\text{observation up to time } t)$ of the signal x_t given the observations y_t for the Ito model:

$$\Sigma_1^{an} : \begin{cases} dx_t &= f(x_t)dt + g(x)dv_t, \quad x(t_0) = x_0 \\ dy_t &= h(x_t)dt + dw_t, \end{cases} \tag{1.76}$$

where $x_t \in \Re^n$, with x_0 a random vector, is a continuous-time process, $\{v_t, t \geq t_0\}$ is an r-dimensional standard Brownian motion, $f : \Re^n \to \Re^n$, $g : \Re \to \Re^{n\times r}$ is the diffusion coefficient, $y_t \in \Re^m$, $h : \Re^n \to \Re^m$ is a known observation function, and $\{w_t, t \geq t_0\}$ is an m-dimensional standard Brownian motion which is independent of v_t and the initial state x_0. Let $\mathcal{F}_t^y = \mathbf{B}(y_s, s \leq t)$ be the filtration produced by the observation process y_t, where $\mathbf{B}(.)$ is the completion of the smallest σ-algebra generated by y_t. With the correction supplied by Kushner (Kushner, 1964a), Stratonovich obtained the following equation

$$dp(x_t|\mathcal{F}_t^y) = \widetilde{A}p(x_t|\mathcal{F}_t^y) + (h - \hat{h})R^{-1}(dz_t), \tag{1.77}$$

with

$$\widetilde{A} = -\frac{\partial}{\partial x}f + \frac{1}{2}Tr\left(gQg^T\frac{\partial^2}{\partial x \partial x}\right),$$

$$\mathsf{E}\{v_t v_\tau^T\} = Q(\tau)\delta_{t,\tau}, \quad \mathsf{E}\{w_t w_\tau^T\} = R(\tau)\delta_{t,\tau}, \quad \hat{h} = \mathsf{E}\{h(x_t|\mathcal{F}_t^y)\},$$

and z_t is the innovation process satisfying

$$dz_t = (h - \hat{h})dt + dw_t. \tag{1.78}$$

Later, Kushner (Kushner, 1964a), (Kushner, 1964b) in 1964 presented his own solution to the filtering problem and obtained an equation for the conditional density and conditional expectation of the observation. Consider a more general nonlinear possibly time-varying model of the form

$$\Sigma_2^{an} : \begin{cases} dx_t &= f(x_t,t)dt + g(x_t,t)dv_t \quad x(t_0) = x_0 \\ dy_t &= h(x_t,t)dt + \sqrt{R_t}dw_t, \end{cases} \tag{1.79}$$

where $f : \Re^n \to \Re^n$,$g : \Re \to \Re^{n\times r}$, $h : \Re^n \to \Re^m$. Then, he obtained an equation for the conditional density and conditional expectation of the observation given by

$$dp(x_t|\mathcal{F}_t^y) = \mathcal{L}p(x_t|\mathcal{F}_t^y)dt + p(x_t|\mathcal{F}_t^y)R_t^{-1}(dy_t - \widehat{h}_t dt), \tag{1.80}$$

where

$$\mathcal{L} = -\frac{\partial}{\partial x}f + \frac{1}{2}Tr\left(gg^T\frac{\partial^2}{\partial x\partial x}\right), \quad \widehat{h}_t = \mathsf{E}\{h(x_t,t)|\mathcal{F}_t^y)\}.$$

Similarly, M. Zakai (Wong, 1965), (Zakai, 1969) in 1969 presented a simpler equation than Kushner's in terms of the unnormalized conditional density $P(x_t|\mathcal{F}_t^y)$ (and linear in it) given by:

$$dP(x_t|\mathcal{F}_t^y) = \mathcal{L}P(x_t|\mathcal{F}_t^y)dt + R_t^{-1}h(x_t,t)P(x_t|\mathcal{F}_t^y)dy_t. \tag{1.81}$$

However, only for the linear Gaussian case and certain class of nonlinearities, can the Zakai equation be solved explicitly. Most of the efforts in this direction have gone into developing numerical schemes for solving both the Kushner and the Zakai equations, which hitherto are neither recursive nor computationally efficient.

1.3.1 Extended Kalman Filters (EKFs) and Unscented Kalman Filters (UKF)

The Kalman filter theory applies to linear-Gaussian problems, but most real-life applications are nonlinear and/or non-Gaussian. Therefore, following the Kalman-Bucy pioneering work, a nonlinear version of the Kalman filter was also developed (the "Extended Kalman Filter" (EKF)) and was actually first derived by Peter Swirling in 1958. The EKF simply approximates the nonlinear model by its first-order Taylor series evaluated at the current estimate. For the system model

$$\Sigma_3^{an} : \begin{cases} \dot{x}(t) &= f(x,t) + G(t)v(t), \quad x(t_0) = x_0 \\ y(t) &= h(x,t) + w(t) \end{cases} \tag{1.82}$$

where $x \in \Re^n$ is the state vector, $y \in \Re^m$ is the measurement or observation vector, $w \in \Re^m$, $v \in \mathbb{R}^r$ are Gaussian noise processes with

$$\mathsf{E}\{v(t)v^T(\tau)\} = R(t)\delta(t-\tau), \quad \mathsf{E}\{w(t)w^T(\tau)\} = Q(t)\delta(t-\tau),$$
$$\mathsf{E}\{v(t)w(\tau)\} = 0 \quad \text{for all } \ t,\tau \in [0,\infty),$$

$f : \Re^n \times \Re \to \Re^n$, $G : \mathbb{R} \to \Re^{n\times r}$, $h : \Re^n \times \Re \to \Re^m$. The EKF is given by

$$\dot{\hat{x}}(t) = f(\hat{x}(t),t) + K(t)[y(t) - h(\hat{x}(t))], \quad \hat{x}(t_0) = \mathsf{E}\{x_0\} \tag{1.83}$$

where

$$K(t) = P(t)H^T(\hat{x}(t),t)R^{-1}(t)$$

and

$$\begin{aligned} \dot{P}(t) &= F(\hat{x}(t),t)P(t) + P(t)F^T(\hat{x}(t),t) - \\ &\quad P(t)H^T(\hat{x}(t),t)R^{-1}(t)H(\hat{x}(t),t)P(t) + G(t)Q(t)G^T(t), \end{aligned}$$

$$P_0 = \mathsf{E}\{x_0x_0^T\}, \quad F(\hat{x}(t),t) = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}}, \quad H(\hat{x}(t),t) = \left.\frac{\partial h}{\partial x}\right|_{\hat{x}}.$$

The EKF, which is suboptimal, has been successfully applied to numerous nonlinear estimation problems (Sorenson, 1985). However, for highly nonlinear problems with large initial errors, divergence may occur. Consequently, the filter must always be initialized in a sufficiently close neighborhood of the initial estimate in order to guarantee convergence.

Furthermore, different EKFs have been derived using various approaches (Daum 2005) such as, (i) different coordinate systems; (ii) different factorization of the covariance matrix; and

(iii) second-order or higher-order Taylor series corrections to the state vector prediction and/or measurement update, etc. Nevertheless, and although the EKFs are widely used, their inaccuracies and limitations have been recognized in the tracking and control communities. Indeed, there is a general consensus that they are: (i) difficult to implement; (ii) difficult to tune; and (iii) only reliable for systems that are almost linear on the time scale of the update interval (Julier, 2000). Consequently, a significant improvement to these EKFs came about with the development of the unscented Kalman-filter (UKF) (Julier, 2000). While the EKFs use a simple linear or first-order Taylor approximation, the UKF uses a more accurate approximation, "*called the unscented transform*" (to evaluate the multidimensional distribution integrals). We summarize the main results of the UKF here. We consider the following nonlinear discrete-time system model

$$\Sigma_1^{dn} : \begin{cases} x_{k+1} &= f(x(k), u(k), v(k), k), \quad x(0) = x_0 \\ y_k &= h(x(k), u(k), k) + w(k) \end{cases} \tag{1.84}$$

where $x(k) \in \Re^n$ is the state vector, $u(k)$ is the input vector, $v(k)$ is the system noise vector, while $w(k)$ is the measurement noise vector, and $y(k) \in \Re^m$ is the observation vector each at time-step k. The noise vectors $v(k)$, $w(k)$ are assumed to have

$$\mathsf{E}\{v(i)v^T(j)\} = Q(i)\delta_{ij}, \quad \mathsf{E}\{w(i)w^T(j)\} = R(i)\delta_{ij}, \quad \mathsf{E}\{v(i)w(j)\} = 0 \text{ for all } i, j. \tag{1.85}$$

The objective is to find the minimum mean-squared error (MMSE) estimate of the state vector conditioned on the observations, or the *conditional mean*, i.e., $\hat{x}(i|j)$ which is given by

$$\hat{x}(i|j) = \mathsf{E}[x_i|Y^j],$$

where $Y^j = \{y(1), \ldots, y(j)\}$. The covariance of the estimate is also denoted by $P(i|j)$. The UKF approximates a nonlinear function by generating a set of points whose sample mean and sample covariance are $\hat{x}(k|k)$, $P(k|k)$ respectively. The nonlinear function is then applied to each of these points inturn to yield a transformed sample. Finally, the predicted mean and covariance are calculated from the transformed sample.

The n-dimensional random state vector $x(k)$ with mean $\hat{x}(k|k)$ and covariance $P(k|k)$ is approximated by $2n+1$ weighted samples or *sigma points* selected by the algorithm

$$S : \begin{cases} \mathcal{X}_0(k|k) &= \hat{x}(k|k) \\ W_0 &= \kappa/n+\kappa) \\ \mathcal{X}_i(k|k) &= \hat{x}(k|k) + \left(\sqrt{(n+\kappa)P(k|k)}\right) \\ W_i &= 1/2(n+\kappa) \\ \mathcal{X}_{i+n}(k|k) &= \hat{x}(k|k) - \left(\sqrt{(n+\kappa)P(k|k)}\right) \\ W_{i+n} &= 1/2(n+\kappa). \end{cases} \tag{1.86}$$

where κ is a real number, $\left(\sqrt{(n+\kappa)P(k|k)}\right)$ is the i-th row or i-th column[1] of the matrix square-root of $(n+\kappa)P(k|k)$, and W_i is the weight that is associated with the i-th point. It can then be proven that, the set of samples S chosen by (1.86) have the same sample mean, covariance and all higher odd-ordered central moments as the distribution of $x(k)$. In addition, the matrix square-root and κ affect the fourth and higher-order moments of the sigma points.

Given the set of samples S generated by (1.86), the prediction steps are as follows.

Algorithm

1. Each sigma point is applied to the process model to obtain the transformed samples

$$\mathcal{X}_i(k+1|k) = f(\mathcal{X}_i(k|k), u(k), k).$$

2. The predicted mean is computed as

$$\hat{x}(k+1|k) = \sum_{i=0}^{2n} W_i \mathcal{X}_i(k+1|k).$$

3. The predicted covariance is similarly computed as

$$P(k+1|k) = \sum_{i=0}^{2n} W_i \{\mathcal{X}_i(k+1|k) - \hat{x}(k+1|k)\}\mathcal{X}_i(k+1|k) - \hat{x}(k+1|k)\}^T.$$

[1]If P is of the form $P = A^T A$, then the sigma points are formed from the rows of A. On the other hand, if $P = AA^T$, then the sigma points are formed from the columns of A.

Consequently, the computations of the mean and covariance in the UKF involve only vector and matrix operations, and does not involve the computation of the Jacobian as in the EKF. It also yields more accurate predictions than those of the EKFs.

1.3.2 Maximum Likelihood Recursive Nonlinear Filtering

In 1968 Mortensen R. E. considered a variational approach to the nonlinear filtering problem using a maximum likelihood function. For the model

$$\Sigma_4^{an} : \begin{cases} \dot{x}(t) &= f(x,t) + v(t) \\ y(t) &= h(x,t) + w(t), \end{cases} \tag{1.87}$$

where all the variables have their previous meanings and dimensions, he obtained the following filter

$$\begin{aligned} \dot{\hat{x}}(t) &= f(\hat{x}(t),t) + \Pi^{-1}(\hat{x},t;\mu,t_0)h_x(\hat{x}(t),t)Q^{-1}(t)[y(t) - h(\hat{x}(t),t)]; \\ & \hat{x}(t_0) = \mu = \mathsf{E}\{x_0\}, \end{aligned} \tag{1.88}$$

where $\Pi^{-1}(t,t_0)$ satisfies a matrix Riccati differential equation with $\Pi^{-1}(t_0,t_0) = \Lambda$ (known) for the suboptimal solution. The optimal solution involves the solution of the following Hamilton-Jacobi-Bellman equation

$$\frac{\partial V}{\partial t}(x,t;\mu;t_0) + H^\star[x, \nabla_x V(x,t;\mu,t_0),t] = 0, \quad V(x,t_0;\mu;t_0) = \frac{1}{2}(x-\mu)^T \Lambda^{-1}(x-\mu) \tag{1.89}$$

and

$$H^\star(x,p,\tau) = -\frac{1}{2}p^T R(\tau)p + p^T f(x,\tau) + \frac{1}{2}[y(\tau) - h(x,\tau)]^T Q^{-1}[y(\tau) - h(x,\tau)].$$

It is possible to show that $\Pi^{-1}(\hat{x},t;\mu,t_0)$ satisfies a matrix Riccati differential equation by computing the total time derivative of $(\Pi)_{ij} = \partial^2 V(x,t;\mu,t_0)/\partial x_i \partial x_j$. However, in addition, one needs to know the components

$$[\partial^2 V(x,t;\mu,t_0)/\partial x_i \partial x_j \partial x_k]\big|_{x=\hat{x}(t)} \tag{1.90}$$

of a tensor of rank 3 along the trajectory $\hat{x}(t)$. Thus, computing the optimal solution is indeed a difficult task.

An alternative possible approximation is to assume that the quantities in (1.90) and the higher derivatives of $V(x,t;\mu,t_0)$ evaluated along $\hat{x}$ are negligible. This is equivalent to assuming that V is quadratic with kernel matrix $\Pi(x,t;\mu,t_0)$. This then leads to the matrix Riccati differential equation for $\Pi^{-1}(t,t_0)$, with appropriate initial condition $\Pi^{-1}(t_0,t_0) = \Lambda$.

1.3.3 Bayesian Nonlinear Filtering and Particle Filters (PFs)

Bayesian methods provide a rigorous general framework for dynamic state estimation problems. The Baysian approach is to construct the PDF of the state based on all the available information. Bayesian nonlinear estimation (Ho, 1964) follows exactly the same procedure as the linear theory outlined in the previous section. Given a set of measurements $\{y_1,\ldots,y_k\}$ of an observation function

$$y = h(x,v), \tag{1.91}$$

where $y \in \Re^m$ is related to the variable of interest $x \in \Re^n$, that is corrupted by a noise process $v \in \Re^r$. Suppose also the joint density function $p(x,v)$ is assumed to be known. The problem is then to find a best estimate $\hat{x}$ of x from this data.

Using the joint density $p(x,v)$, the marginal densities $p(x)$ and $p(v)$ can be readily obtained. Then the Bayesian solution can be determined in the following steps.

1. Evaluate $p(y)$: this can be achieved analytically (in principle) or experimentally by Monte-Carlo methods.

2. Evaluate either (a) $p(x,y)$ or (b) $p(x/y)$ in the following way:

(a) $p(x, y)$ can be obtained analytically if v is of the same dimension as y and one can obtain the functional relationship $v = h^{\star}(x, y)$ from (1.91). Then using $p(x, y)$ we have

$$p(x, y) = p(x, v = h^{\star}(x, y)) = h^{\star}(x, y)\mathsf{J}$$

where J is the Jacobian matrix $\mathsf{J} = \det\left(\frac{\partial h^{\star}(x,y)}{\partial y}\right)$.

(b) $p(x/y)$ can also be obtained either analytically or experimentally from $y = h(x, v)$ and $p(x, v)$.

3. Evaluate $p(x/y)$ using either of the following relationships:

(a) Following step 2(a) above,

$$p(x/y) = \frac{p(x, y)}{p(y)}.$$

(b) Following step 2(b) above, one uses Bayes' rule to get

$$p(x/y) = \frac{p(y/x)p(x)}{p(y)}.$$

The above step may be easy or difficult depending on the distribution one has assumed for or obtained for $p(x, v)$, $p(y)$, $p(y/x)$.

4. The *"a posteriori"* density function $p(x/y)$ contains all the information necessary for the estimation of x. One can use several criterion functions for estimating $\hat{x}$ from $p(x/y)$:

(a) Maximize $J_1 = \{prob(\hat{x} = x)\}$. The solution is $\hat{x} =$ Mode of $p(x/y)$, and is also known as the most probable estimate. When the a priori density function $p(x)$ is uniform, this estimate coincides with the classical maximum likelihood estimate.

(b) Minimize $J_2 = \int \|x - \hat{x}\|^2 p(x/y)dx$. The solution is $\hat{x} = E\{x/y\}$ and is known as the conditional mean estimate.

(c) Minmax $J_3 = |x - \hat{x}|$. The solution is $\hat{x} =$ Median of $p(x/y)$, and is known as the minimax estimate.

Other criterion functions could also be used. References (Arasaratnam, 2007a), (Arasaratnam, 2009) also present latest approaches to the Bayesian approach for discrete-time nonlinear systems.

However, an important class of Bayesian filters that were developed as improvements over the EKF are called "*particle filters (PFs)*" (Gordon, 1993). For linear Gaussian estimation, the required PDF remains Gaussian at every iteration of the filter. However, for nonlinear or non-Gaussian problems, there is no general analytic (closed-form) expression for the required PDF. Thus, the central idea behind PFs is to represent the required PDF as a set of random samples, rather than as a function over the state-space. Moreover, as the number of samples become large, they effectively provide an exact equivalent representation of the required PDF. Similarly, estimates of the moments (such as mean and covariance) of the state vector PDF and its functional representation can be obtained or constructed directly from the samples.

A recursive weighted bootstrap algorithm which is based on Baye's rule is used to update the samples. The samples are naturally concentrated in the regions of high probability density. They also have the great advantage of being able to handle any functional nonlinearity, as well as system and measurement noise of any distribution. To review the approach, we first consider the following general discrete-time nonlinear system

$$\Sigma_3^{dn} : \begin{cases} x_{k+1} &= f_k(x_k, w_k), \quad x(0) = x_0 \\ y_k &= h_k(x_k, v_k) \end{cases} \tag{1.92}$$

where $f_k : \Re^n \times \Re^m \to \Re^n, k = 1, \ldots,$ is the system transition function, $h_k : \Re^n \times \Re^r, k = 1, \ldots,$ is the measurement function, and w_k, v_k are uncorrelated zero-mean white noise sequences of known PDF. It is assumed that the initial PDF of the state vectors $p(x_1|D_0) \equiv p(x_1)$ is available, where $D_k = \{y_0, \ldots, y_k\}$ is the measurement information at time k.

The objective is to construct the PDF of the current state x_k given all the available information, $p(x_k|D_k)$. This can theoretically be obtained in two steps: a prediction step, and an update step. For if we suppose $p(x_{k-1}|D_{k-1})$ is available at step $k-1$. Then $p(x_k|D_{k-1})$

can be obtained as

$$p(x_k|D_{k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|D_{k-1})dx_{k-1} \quad \text{—Prediction equation} \tag{1.93}$$

The state transition probabilities $p(x_k|x_{k-1})$, which are assumed to be Markovian, are defined by the systems equations and the known statistics of w_{k-1}

$$p(x_k|x_{k-1}) = \int p(x_k|x_{k-1}, w_{k-1})p(w_{k-1}|x_{k-1})dw_{k-1}, \tag{1.94}$$

and since by assumption $p(w_k|x_{k-1}) = p(w_{k-1})$, we have

$$p(x_k|x_{k-1}) = \int \delta(x_k - f_{k-1}(x_{k-1}, w_{k-1}))p(w_{k-1})dw_{k-1}. \tag{1.95}$$

If now at time step k, a measurement y_k becomes available, then (1.93) can be updated as

$$p(x_k|D_k) = \frac{p(y_k|x_k)p(x_k|D_{k-1})}{\int p(y_k|x_k)p(x_k|D_{k-1})dx_k} \quad \text{—Update equation,} \tag{1.96}$$

where again the conditional PDF $p(y_k|x_k)$ is defined by the measurement model and the known statistics of v_k,

$$p(y_k|x_k) = \int \delta(y_k - h_k(x_k, v_k))p(v_k)dv_k. \tag{1.97}$$

The above steps summarize the theoretical Bayesian estimation algorithm. However, analytical solutions to this problem are only available for a relatively small and restrictive choice of systems and measurement models, e.g. the Kalman filter, where f_k and h_k are linear, while w_k and v_k are additive Gaussian of known variance. In reality, these assumptions are unreasonable for many applications, and hence the need to modify the approach to conform with more realistic situations.

Consequently, the "bootstrap filter" is developed to address some of the above concerns. Suppose a set of random samples $\{x_{k-1}(i) : i = 1, \ldots, N\}$ from the PDF $p(x_{k-1}|D_{k-1})$ are available. The algorithm propagates and updates these samples to obtain a set of new values

$\{x_k(i) : i = 1, \ldots, N\}$, which are approximately distributed as $p(x_k|D_k)$.

Bootstrap Filter Algorithm:

- *Prediction:* The system state equations (1.92) is applied on each sample to obtain new values as

$$x_k^\star(i) = f_{k-1}(x_{k-1}(i), w_{k-1}(i)), \quad i = 1, \ldots, N,$$

where $w_{k-1}(i), i = 1, \ldots, N$ is drawn from the assumed PDF of w_{k-1}.

- *Update:* On receipt of a new measurement y_k, evaluate the likelihood of each prior sample to obtain a normalized weight for each sample

$$q_i = \frac{p(y_k|x_k^\star(i))}{\sum_{j=1}^{N} p(y_k|x_k^\star(j))}, \quad i = 1, \ldots, N.$$

In this way, a discrete distribution over $\{x_k^\star(i) : i = 1, \ldots, N\}$ with probability mass q_i associated with each sample i is defined. Next, resample N times from the discrete distribution to generate samples $\{x_k(i) : i = 1, \ldots, N\}$ so that for any j, $Pr\{x_k(j) = x_k^\star(i)\} = q_i$.

The above steps form a single iteration of the recursive filter algorithm. To initialize the algorithm, N samples $x_k^\star(i)$ are drawn from the known prior $p(x_1)$, and are then applied directly to the update step of the algorithm. The claim is that the samples $x_k(i)$ are approximately distributed as the required PDF of $p(x_k|D_k)$.

The above basic algorithm is simple and easy to program. The only requirements are

(a) $p(x_1)$ is available for sampling;

(b) $p(y_k|x_k)$ is a known functional form;

(c) $p(w_k)$ is available for sampling.

The output of the algorithm as a set of samples of the required posterior density is also convenient for many applications. In addition, it is straightforward to obtain estimates of the mean and covariance of the state, and indeed any function of the state.

1.4 Conclusion

In this chapter, we have reviewed the historical development of estimation theory from Gauss's least squares method to the Kalman-Bucy theory and finally the Stratonovich-Kushner theory. We have summarized most of the major approaches that have been developed for linear dynamic systems, including the minimum-variance method, the maximum likelihood method and the Bayesian approaches. Finally, we have also discussed the extensions of the above approaches to nonlinear dynamic systems including the extended Kalman filter (EKF), the Stratonovich and Kushner filters, as well as the maximum likelihood recursive nonlinear filters and Bayesian nonlinear filters. In the next chapter, we focus on linear singular systems.

CHAPTER 2

LITERATURE REVIEW

Singular systems are classified into two types; namely, (i) singularly-perturbed systems; and (ii) differential-algebraic systems, descriptor or generalized state-space systems. The two groups are also related as the second group can be obtained from the first by asymptotically allowing the pertubation parameter to go zero.

Singularly perturbed systems are that class of systems that are characterized by a discontinuous dependence of the system properties on a small perturbation parameter ϵ. They arise in many physical systems such as electrical power systems and electrical machines (e.g. an asynchronous generator, a dc motor, electrical converters), electronic systems (e.g. oscillators) mechanical systems (e.g. fighter aircrafts), bilogical systems (eg. bacterial-yeast-virus cultures, heart) and also economic systems with various competing sectors. This class of systems has two time-scales; namely, a "fast" and a "slow" dynamics. This makes their analysis and control more complicated than regular systems. Nevertheless, they have been studied extensively (Khalil, 1985), (Kokotovic, 1986).

The filtering problem for linear singularly perturbed systems in both continuous-time (Assawinchaichote, 2004a)-(Assawinchaichote, 2007), (Gajic, 1994), (Haddad, 1976)-(Hong, 2008), (Lim, 2000), (Mukaidani, 2003), (Prljaca, 2008), (Sebald, 1978), (Shen, 1993), (Shen, 1996), (Yang, 2008) and discrete-time (Kim 2002), (Lim, 1996), (Sadjadi, 1990) has been considered by many authors. Various types of filters have been proposed, including composite (Haddad, 1976), (Haddad, 1977), (Sebald, 1978), (Shen, 1993) and reduced-order filters (Gajic, 1994), (Sebald, 1978), (Shen, 1993).

On the other hand, descriptor, differential or generalized state-space systems provide a more generalized description of dynamic systems including possible constraints conditions on the states and the effect of small parameter perturbation (or singular-perturbation) in the model.

They are also encountered in chemical and minerals industries, mechanical and aerospace systems, as well as electronic and electrical circuits. Because of the incorporation of the constraints conditions in the state equations, the state-variables are usually referred to as semistate variables (Newcomb, 1981a).

Similarly, various authors have considered the observer design and filtering problems for linear descriptor systems in both continuous-time (Dai, 1989), (Dai 1989), (Darouach, 1995)-(Darouach, 1997), (El-Tohami, 1984), (Fahmy, 1989), (Gao, 2004), (Hou, 1995), (Ishihara, 2009), (Koenig, 1995), (Minamide, 1989), (Paraskevopoulos, 1992), (Sun, 2007), (Uetake, 1989), (Zhou, 2008) and discrete-time (Dai 1988), (Darouach, 2009), (Boulkroune, 2010), (El-Tohami, 1987)-(El-Tohami, 1983), (Ishihara, 2006), (Nikoukhah, 1999), (Nikoukhah, 1992), (Zhou, 2008). Kalman - Luenberger type full-order and reduced-order observers have extensively been studied, and necessary and sufficient conditions for the solvability of the problem have been presented. On the other hand, only recently has there been some attention on the design of observers and filters for nonlinear descriptor systems (Darouach, 2008). This is probably because of the complexity of this class of systems. Similarly, the output observation design problem for nonlinear systems has also been considered in (Zimmer, 1997). But to the best of our knowledge, the filtering problem for more general affine nonlinear descriptor systems has not been discussed in any reference.

In this chapter, we review some of the methods for both the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ filtering problems and for both linear singularly perturbed and linear descriptor systems respectively. But only the continuous time results will be presented. The chapter is organized as follows. In section 2.1 we discuss the $\mathcal{H}_2$ filtering problem for the linear singularly perturbed case. While in Section 2.2, we discuss the $\mathcal{H}_\infty$ problem. Then in Sections 2.3 and 2.4, we discuss the corresponding linear descriptor problems respectively.

2.1 Review of Kalman ($\mathcal{H}_2$)-Filtering for Linear Singularly-Perturbed Systems

In this section, we review Kalman (or $\mathcal{H}_2$) filtering results for linear singularly-perturbed systems and in the subsequent section, we consider the $\mathcal{H}_\infty$ problem. The results presented here are mainly from (Haddad, 1976).

We consider the following linear (possibly time-varying) (LTV) singularly-perturbed system

$$\Sigma^l_{1,\mu} : \begin{cases} \dot{x}_1 &= A_1(t)x_1 + A_{12}(t)x_2 + B_1(t)w, \quad x_1(t_0,\epsilon) = x_{10} \\ \mu\dot{x}_2 &= A_{21}(t)x_1 + A_2x_2 + B_2(t)w, \quad x_2(t_0,\epsilon) = x_{20} \\ y &= C_1(t)x_1 + C_2(t)x_2 + v \end{cases} \tag{2.1}$$

where $x_1 \in \Re^{n_1}$ is the slow state vector, $x_2 \in \Re^{n_2}$ is the fast state vector, while $y \in \Re^m$ is the output measurement vector. The vectors w, v are uncorrelated white noise processes with covariances given by

$$\mathsf{E}\{w(t)w^T(\tau)\} = Q(t)\delta(t-\tau), \quad \mathsf{E}\{v(t)v^T(\tau)\} = R(t)\delta(t-\tau),$$

while the matrices $A_1(t)$, $A_{21}(t)$, $A_{12}(t)$, $A_2(t)$ are continuous with respect to t and have appropriate dimensions, and μ is a small parameter. It is also assumed that $A_2(t)$ is stable, nonsingular, and has bounded first derivative.

Applying the Chang (Chang, 1972) transformation, the above system (2.1) can be transformed to the following decomposed system

$$\widetilde{\Sigma}^l_{2,\mu} : \begin{cases} \dot{\eta} &= A_0(t)\eta + B_0(t)w, \quad \eta(t_0,\epsilon) = \eta_0 \\ \mu\dot{\xi} &= A_2(t)\xi + B_2(t)w, \quad \xi(t_0,\epsilon) = \xi_0 \\ y &= C_0(t)\eta_0 + C_2(t)\xi + v, \quad t \geq t_0 \end{cases} \tag{2.2}$$

where again $\eta \in \Re^{n_1}$ is the slow state vector, $\xi \in \Re^{n_2}$ is the fast state vector, while all other variables retain their previous meanings,

$$A_0(t) = A_1(t) - A_{12}(t)\mathsf{L}(t) + O(\mu)$$

$$
\begin{aligned}
B_0(t) &= B_1(t) - \mu H(t)\mathsf{L}(t)B_1(t) - \mathsf{H}(t)B_2(t) = B_1(t) - A_{12}(t)A_2^{-1}(t)B_2 + O(\mu) \\
C_0(t) &= C_1(t) - C_2(t)\mathsf{L}(t) = C_1(t) - C_2(t)A_2^{-1}(t)A_{21}(t) + O(\mu) \\
\mu\dot{\mathsf{L}}(t) &= A_2(t)\mathsf{L}(t) - A_{21}(t) - \mu\mathsf{L}(t)(A_1(t) - A_{12}(t)\mathsf{L}(t)) \\
\mu\dot{\mathsf{H}}(t) &= -\mathsf{H}(t)(A_2(t) + \mu\mathsf{L}(t)A_{12}(t)) + A_{12}(t) + \mu(A_1(t) - A_{12}(t)\mathsf{L}(t))\mathsf{H}(t)
\end{aligned}
$$

and all matrices have appropriate dimensions. The problem is to find the best estimates $\hat{\eta}$, $\hat{\xi}$ of η, ξ from the measurements $y(t)$, that minimize the mean-squared errors $\mathsf{E}\{\|\eta(t)-\hat{\eta}(t)\|^2\}$, $\mathsf{E}\{\|\xi(t)-\hat{\xi}(t)\|^2\}$ and to investigate the behavior of the resulting filters as $\mu \to 0$.

It can be shown that (Haddad, 1976) the solution to the above problem is given by the following filter

$$
\mathbf{F}^l_{1,\mu} : \begin{cases}
\dot{\hat{\eta}}(t) = A_0(t)\hat{\eta} + (P_1(t)C_0^T(t) + P_{12}(t)C_2^T(t))R^{-1}(t)(y - C_0(t)\hat{\eta}(t) - C_2\hat{\xi}(t)); \\
\qquad \hat{\eta}(t_0) = \mathsf{E}\{\eta_0\} \\
\mu\dot{\hat{\xi}}(t) = A_2(t)\hat{\xi}(t) + (\mu P_{12}(t)C_0^T(t) + P_2(t)C_2^T(t))R^{-1}(t)(y(t) - C_0(t)\hat{\eta}(t) - \\
\qquad C_2(t)\hat{\xi}(t)), \quad \hat{\xi}(t_0) = \mathsf{E}\{\eta_0\}
\end{cases}
\tag{2.3}
$$

where

$$
\begin{aligned}
\dot{P}_1(t) &= A_0(t)P_1(t) + P_1(t)A_0^T(t) + B_0(t)Q(t)B_0^T(t) - (P_1(t)C_0^T(t) + \\
&\quad P_{12}(t)C_0^T(t))R^{-1}(t)(C_0(t)P_1(t) + C_2(t)P_{12}(t)), \;\; P_1(t_0) = Cov\{\eta_0\} \qquad (2.4) \\
\mu\dot{P}_{12}(t) &= \mu A_0(t)P_{12}(t) + P_{12}(t)A_2^T(t) + B_0(t)Q(t)B_0^T(t) - (P_1(t)C_0^T(t) + \\
&\quad P_{12}(t)C_2^T(t))R^{-1}(t)(\mu C_0(t)P_{12}(t) + C_2(t)P_2(t)), \; P_{12}(t_0) = Cov\{\eta_0, \xi_0\} \quad (2.5)
\end{aligned}
$$

and

$$
P_1(t) = \mathsf{E}\{\|\eta(t) - \hat{\eta}(t)\|^2\}, \;\; P_2(t)/\mu = \mathsf{E}\{\|\xi(t) - \hat{\xi}(t)\|^2\}, \;\; P_2(t_0) = \mu Cov\{\xi_0\}. \tag{2.6}
$$

The limiting behaviors of the above Riccati equations (2.4)-(2.6) as $\mu \to 0$ can be obtained by expressing each of the matrices $P_2(t)$ and $P_{12}(t)$ as a sum of a steady-state term and a boundary layer term up to order $O(\mu)$ as

$$P_2(t) = \bar{P}_2(t) + \tilde{P}_2(t) + O(\mu), \quad t \geq t_0 \tag{2.7}$$

$$P_{12}(t) = \bar{P}_{12}(t) + \tilde{P}_{12}(t) + O(\mu), \quad t \geq t_0 \tag{2.8}$$

respectively. The steady-state terms $\bar{P}_2(t)$, $\bar{P}_{12}(t)$ are obtained by setting $\mu = 0$ in (2.5) to get

$$A_2(t)\bar{P}_2(t) + \bar{P}_2(t)A_2^T(t) + B_2(t)Q(t)B_2^T(t) - \bar{P}_2(t)C_2^T(t)R^{-1}(t)C_2(t)\bar{P}_2(t) = 0, \tag{2.9}$$

$$\bar{P}_{12}(t) = -[B_0(t)Q(t)B_2^T(t) - P_1(t)C_0^T(t)R^{-1}(t)C_2(t)\bar{P}_2(t)]\bar{A}_2^{-1}(t), \tag{2.10}$$

where

$$\bar{A}_2(t) = A_2(t) - \bar{P}_2(t)C_2^T(t)R^{-1}(t)C_2(t),$$

and $\bar{P}_2(t)$ is chosen as the positive semidefinite solution of (2.9). Whereas the boundary-layer terms $\tilde{P}_2(t)$, $\tilde{P}_{12}(t)$ are obtained as the solutions to the following differential equations in the stretched time variable $\tau = (t - t_0)/\mu$ as

$$\begin{aligned} \frac{d}{d\tau}\tilde{P}_2(\tau) &= \bar{A}_2(t_0)\tilde{P}_2(\tau) + \tilde{P}_2(\tau)\bar{A}_2^T(t_0) - \tilde{P}_2(\tau)C_2^T(t_0)R^{-1}C_2(t_0)\tilde{P}_2(\tau) + O(\mu), \\ &\quad \tilde{P}_2(0) = \mu Cov\{\xi_0\} - \bar{P}_2(t_0), \ \ \tau \geq 0, \end{aligned} \tag{2.11}$$

$$\begin{aligned} \frac{d}{d\tau}\tilde{P}_{12}(\tau) &= \tilde{P}_{12}(\tau)[\bar{A}_2^T(t_0) - C_2^T(t)R^{-1}(t)C_2(t)\tilde{P}_2(\tau)] - [P_1(t_0)C_0^T(t) + \\ &\quad \tilde{P}_{12}(t_0)C_2^T(t)]R^{-1}C_2(t)\tilde{P}_2(\tau), \ \tilde{P}_{12}(\tau) = Cov\{\eta_0, \xi_0) - \bar{P}_{12}(t_0), \ \tau \geq 0 \end{aligned} \tag{2.12}$$

Since $A_2(t)$ is stable and $\bar{P}_2 \geq 0$, it implies that $\bar{A}_2(t)$ is also a stable matrix. Consequently, both $\tilde{P}_2(\tau)$ and $\tilde{P}_{12}(\tau)$ tend to zero as $\tau \to \infty$ and $\mu \to 0$. The limiting behavior of $P_1(t)$ also follows from (2.4) as

$$\begin{aligned} \dot{P}_1(t) &= A_0(t)P_1(t) + P_1(t)A_0^T(t) + B_0(t)Q(t)B_0^T(t) - (P_1(t)C_0^T(t) + \\ &\quad \bar{P}_{12}(t)C_0^T(t))R^{-1}(t)(C_0(t)P_1(t) + C_2(t)\bar{P}_{12}(t)) + O(\mu), \\ &\quad P_1(t_0) = Cov\{\eta_0\}, \ \ t \geq t_0. \end{aligned} \tag{2.13}$$

The limiting behaviors of the filters can also be studied. For the fast mode filter in the stretched time parameter τ, the effect of $\tilde{P}_2(t)$ can be neglected and for $t_2 \geq t \geq t_1 \geq t_0 + \epsilon$, $\epsilon > 0$ arbitrary,

$$\frac{d}{d\tau}\hat{\xi} = \bar{A}_2(t_1)\hat{\xi}_2(t) + \bar{P}_2(t_1)C_2^T R^{-1}(y - C_0\hat{\eta}) + O(\mu^{\frac{1}{2}}), \quad \tau \geq 0. \tag{2.14}$$

However, near the initial estimation interval $\epsilon \geq (t - t_0) \geq 0$ the above filter (2.14) needs to be modified by adding $\tilde{P}_2(\tau)$ to $\bar{P}_2(\tau)$.

Similarly, for the slow-mode filter, we have

$$\dot{\hat{\eta}} = A_0(t)\hat{\eta} + K_1(t)(y(t) - C_0(t)\hat{\eta}) + O(\mu^{\frac{1}{2}}), \tag{2.15}$$

where

$$K_1(t) = (P_1(t)C_0^T(t) + B_0(t)Q(t)D_0^T(t))R_0^{-1}(t),$$

with

$$R_0(t) = (R(t) + D_0Q(t)D_0^T(t)), \quad D_0(t) = C_2(t)A_2^{-1}(t)B_2(t).$$

The reduced-order filter can also be obtained by setting $\mu = 0$ in (2.2) to get

$$\bar{\xi} = -A_2^{-1}(t)B_2(t)w,$$

which yields

$$\bar{y}(t) = C_0(t)\eta - C_2(t)A_2^{-1}(t)B_2(t)w + v = C_0(t)\eta + D_0(t)w + v.$$

The reduced filter can then be constructed as

$$\dot{\hat{\eta}} = A_0(t)\eta + K_0(t)[\bar{y} - C_0(t)\eta], \quad \hat{\eta}(t_0) = \mathsf{E}\{\eta_0\}, \tag{2.16}$$

where

$$K_0(t) = (P_0(t)C_0^T(t) + B_0(t)Q(t)D_0^T(t))R_0^{-1},$$

$$R_0(t) = (R(t) + D_0 Q(t) D_0^T),$$

and

$$\begin{aligned}
\dot{P}_0(t) &= A_0(t)P_0(t) + P_0(t)A_0^T(t) + B_0(t)Q(t)B_0^T(t) - \\
&\quad (P_0(t)C_0^T(t) + B_0(t)Q(t)D_0^T(t))R_0^{-1}(t)(C_0(t)P_0(t) + D_0(t)Q(t)B_0^T(t)), \\
&\quad P_0(t_0) = Cov\{\eta_0\}.
\end{aligned} \tag{2.17}$$

Next, we consider the time-invariant case in which the system matrices in (2.1) are constant, i.e.,

$$\Sigma_{3\mu}^l : \begin{cases} \dot{x}_1 = A_1 x_1 + A_{12} x_2 + B_1 w, & x_1(t_0, \epsilon) = x_{10} \\ \mu \dot{x}_2 = A_{21} x_1 + A_2 x_2 + B_2 w, & x_2(t_0, \epsilon) = x_{20} \\ y = C_1 x_1 + C_2 x_2 + v. \end{cases} \tag{2.18}$$

In addition, $x_1(t_0)$, $x_2(t_0)$ are also assumed to be random vectors with mean values $\mathsf{E}\{x_1\} = \bar{x}_1$, $\mathsf{E}\{x_2\} = \bar{x}_2$, and the covariances are also positive-definite constant matrices

$$\mathsf{E}\{w(t)w^T(\tau)\} = Q\delta(t-\tau), \quad \mathsf{E}\{v(t)v^T(\tau)\} = R\delta(t-\tau).$$

A time-invariant aggregate filter can be constructed for the system (Gajic, 1994) as

$$\mathbf{F}_{2,\mu}^l : \begin{cases} \dot{\hat{x}}_1 = A_1 \hat{x}_1 + A_{12} \hat{x}_2 + K_1 e, & \hat{x}_1(t_0) = \bar{x}_{10} \\ \mu \dot{\hat{x}}_2 = A_{21} \hat{x}_1 + A_2 \hat{x}_2 + K_2 e, & \hat{x}_2(t_0) = \bar{x}_{20} \\ e = y - C_1 \hat{x}_1 + C_2 \hat{x}_2, \end{cases} \tag{2.19}$$

where, the optimal filter gains K_1, K_2 are obtained from (Khalil, 1984) as

$$K_1 = (P_1 C_1^T + P_{12} C_2^T) R^{-1}, \tag{2.20}$$

$$K_2 = (\mu P_{12}^T C_1^T + P_2 C_2^T) R^{-1}, \tag{2.21}$$

with the matrices P_1, P_{12}, and P_2 representing the positive semidefinite stabilizing solution of the filter algebraic Riccati equation (ARE):

$$AP + PA^T + PSP + BQB^T = 0, \tag{2.22}$$

where

$$A = \begin{bmatrix} A_1 & A_{12} \\ \frac{1}{\mu}A_{21} & \frac{1}{\mu}A_2 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ \frac{1}{\mu}B_2 \end{bmatrix}, \quad P = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & \frac{1}{\mu}P_2 \end{bmatrix},$$
$$C = [C_1 \quad C_2], \quad S = C^T R^{-1} C.$$

The above filter (2.19) can also be decomposed using the Chang transformation (Gajic, 1994)

$$\begin{bmatrix} \hat{\eta}_1 \\ \hat{\eta}_1 \end{bmatrix} = \begin{bmatrix} I - \mu \mathsf{HL} & -\mu \mathsf{H} \\ \mathsf{L} & I \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}, \tag{2.23}$$

where L and H satisfy the algebraic equations

$$\begin{aligned} A_2\mathsf{L} - A_{21} - \mu\mathsf{L}(A_1 - A_{12}\mathsf{L}) &= 0 \\ -\mathsf{H}A_2 + A_{12} - \mu\mathsf{HL}A_{12} + \mu(A_1 - A_{12}\mathsf{L})\mathsf{H} &= 0 \end{aligned}$$

to get

$$\mathbf{F}_{3,\mu}^l : \begin{cases} \dot{\hat{\eta}}_1 = (A_1 - A_{12}\mathsf{L})\hat{\eta}_1 + (K_1 - \mathsf{H}K_2 - \mu\mathsf{HL}K_1)e, & \hat{\eta}_1(t_0) = \bar{\eta}_{10} \\ \mu\dot{\hat{\eta}}_2 = (A_2 + \mu\mathsf{L}A_{12})\hat{\eta}_2 + (K_2 + \mu\mathsf{L}K_1)e, & \hat{\eta}_2(t_0) = \bar{\eta}_{20} \\ e = y - (C_1 - C_2\mathsf{L})\hat{\eta}_1 - [C_2 + \mu(C_1 - C_2\mathsf{L})\mathsf{H})]\hat{\eta}_2. \end{cases} \tag{2.24}$$

2.2 Review of ($\mathcal{H}_\infty$)-Filtering for Linear Singularly-Perturbed Systems

In this section, we present some results from (Lim, 2000), (Shen, 1996) on the $\mathcal{H}_\infty$ filtering problem for linear singularly-perturbed systems. We reconsider the LTI model of the system (2.18), and where $w, v \in \mathcal{L}_2[t_0, \infty)$. It is desired to design an estimator of the form (2.19) to estimate the states of the system so that the following objective is achieved:

$$\sup_{w,v} J = \sup_{w,v} \frac{\int_{t_0}^{\infty} \|z(t) - \hat{z}(t)\|_R^2}{\int_{t_0}^{\infty} (\|w(t)\|_{W^{-1}}^2 + \|v(t)\|^2) dt} \leq \gamma^2 \tag{2.25}$$

for some weighting matrices $R \geq 0$, $W > 0$, a prescribed number $\gamma > 0$, and where the penalty variable z is a linear combination of the states defined as

$$z = G_1 x_1 + G_2 x_2.$$

As in the $\mathcal{H}_2$ problem presented in the previous section, the aggregate $\mathcal{H}_\infty$ filter solution is given by the following gains

$$K_1 = \widetilde{P}_1 C_1^T + \widetilde{P}_2 C_2^T, \quad K_2 = \mu \widetilde{P}_2^T C_1^T + \widetilde{P}_3 C_2^T \tag{2.26}$$

with the matrices $\widetilde{P}_1$, $\widetilde{P}_2$, and $\widetilde{P}_3$ representing the positive semidefinite stabilizing solution of the filter algebraic-Riccati-equation (ARE):

$$A\widetilde{P} + \widetilde{P}A^T - \widetilde{P}\left(C^T C - \frac{1}{\gamma^2} G^T R G\right)\widetilde{P} + BWB^T = 0, \tag{2.27}$$

where

$$\widetilde{P} = \begin{bmatrix} \widetilde{P}_1 & \widetilde{P}_2 \\ \widetilde{P}_2^T & \frac{1}{\mu}\widetilde{P}_3 \end{bmatrix}, \quad G = [G_1 \quad G_2],$$

and all the other matrices are as defined before.

A more numerically efficient and well-conditioned decomposition filter comprising of a separate pure-slow and pure-fast independent filters directly driven by the innovation signal can also be obtained. Accordingly, consider the optimal closed-loop filter equations (2.19), (2.26)

$$\mathbf{F}_{4,\mu}^l : \begin{cases} \dot{\hat{x}}_1 &= (A_1 - K_1 C_1)\hat{x}_1 + (A_{12} - K_1 C_2)\hat{x}_2 + K_1 y, \\ \mu\dot{\hat{x}}_2 &= (A_{21} - K_2 C_1)\hat{x}_1 + (A_2 - K_2 C2)\hat{x}_2 + K_2 y. \end{cases} \tag{2.28}$$

Then, there exists a nonsingular transformation T and a change of coordinates

$$\begin{bmatrix} \zeta_s \\ \zeta_f \end{bmatrix} = \mathsf{T} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

which transforms the above filter (2.28) into

$$\mathbf{F}^l_{5,\mu} : \left\{ \begin{bmatrix} \dot{\hat{\zeta}}_s \\ \dot{\hat{\zeta}}_f \end{bmatrix} = \begin{bmatrix} a_s & 0 \\ 0 & \frac{1}{\mu} a_f \end{bmatrix} \begin{bmatrix} \hat{\zeta}_s \\ \hat{\zeta}_f \end{bmatrix} + \begin{bmatrix} K_s \\ \frac{1}{\mu} K_f \end{bmatrix} y \right. \tag{2.29}$$

To proceed, we make the folowing assumption.

Assumption 2.2.1. *The triple (A_2, C_2, B_2) is controllable and observable.*

Then, define the following transformation matrices

$$\begin{aligned}
T_1 &= \begin{bmatrix} A_1^T & -\left(C_1^T C_1 - \frac{1}{\gamma^2} G_1^T R G_1\right) \\ -B_1 W B_1^T & -A_1 \end{bmatrix}, \\
T_2 &= \begin{bmatrix} A_{21}^T & -\left(C_1^T C_2 - \frac{1}{\gamma^2} G_1^T R G_2\right) \\ -B_1 W B_2^T & -A_{12} \end{bmatrix}, \\
T_3 &= \begin{bmatrix} A_{12}^T & -\left(C_2^T C_1 - \frac{1}{\gamma^2} G_2^T R G_1\right) \\ -B_2 W B_1^T & -A_{21} \end{bmatrix}, \\
T_4 &= \begin{bmatrix} A_2^T & -\left(C_2^T C_2 - \frac{1}{\gamma^2} G_2^T R G_2\right) \\ -B_2 W B_2^T & -A_2 \end{bmatrix}.
\end{aligned}$$

Then, it can be shown that $T_i, i = 1, \ldots, 4$ are components of the Hamiltonian matrix of the system defined as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{p}_1 \\ \dot{x}_2 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} T_1 & T_2 \\ \frac{1}{\mu} T_3 & \frac{1}{\mu} T_4 \end{bmatrix} \begin{bmatrix} x_1 \\ p_1 \\ x_2 \\ p_2 \end{bmatrix}, \tag{2.30}$$

where p_1, p_2 are the corresponding costate vectors. The slow-fast decomposition of the above aggregate filter can then be achieved by again using the Chang transformation

$$T = \begin{bmatrix} I - \mu NM & -\mu N \\ M & I \end{bmatrix},$$

for some matrices M, N satisfying simultaneously the pair of algebraic equations

$$T_4 M - T_3 - \mu M(T_1 - T_2 M) = 0 \tag{2.31}$$

$$-N(T_4 + \mu M T_2) + T_2 + \mu(T_1 - T_2 M)N = 0. \tag{2.32}$$

Note that Assumption 2.2.1 guarantees that T_4 is nonsingular and there exists a solution to the above algebraic equations. This is guaranteed by the Implicit-function theorem, and can be obtained by using iterative methods, e.g. the Newton's method with initial conditions $M^{(0)} = M + O(\mu) = T_4^{-1} T_3$, $N^{(0)} = N + O(\mu) = T_2 T_4^{-1}$. Then, using similar results as in (Gajic, 1994), the solution of the ARE (2.27) can be related to the solutions of the pure-slow and pure-fast AREs as

$$P = \left(\Omega_3 + \Omega_4 \begin{bmatrix} P_s & 0 \\ 0 & P_f \end{bmatrix} \right) \left(\Omega_1 + \Omega_2 \begin{bmatrix} P_s & 0 \\ 0 & P_f \end{bmatrix} \right)^{-1} \tag{2.33}$$

where P_s, P_f satisfy the AREs

$$P_s a_1 - a_4 P_s - a_3 + P_s a_2 P_s = 0, \tag{2.34}$$

$$P_f b_1 - b_4 P_f - b_3 + p_f b_2 p_f = 0, \tag{2.35}$$

with

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = T_1 - T_2 M, \quad \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = T_4 + \mu M T_2,$$

while the matrices $\Omega_i, i = 1, \ldots, 4$ are given by

$$\begin{bmatrix} \Omega_1 & \Omega_2 \\ \Omega_3 & \Omega_4 \end{bmatrix} = E_1^{-1} T^{-1} E_2 = E_1^{-1} \begin{bmatrix} I & \mu N \\ -M & I - \mu N M \end{bmatrix} E_2,$$

with the permutation matrices E_1, E_2 given by

$$\begin{bmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & 0 & I_{n_1} & 0 \\ 0 & \frac{1}{\mu} I_{n_2} & 0 & 0 \\ 0 & 0 & 0 & I_{n_2} \end{bmatrix}, \quad \begin{bmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & 0 & I_{n_1} & 0 \\ 0 & I_{n_2} & 0 & 0 \\ 0 & 0 & 0 & I_{n_2} \end{bmatrix}.$$

Now, since

$$\begin{aligned} \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} &= T_1 - T_2(M^{(0)} + O(\mu)) = T_1 - T_2 T_4^{-1} T_3 + O(\mu) \\ &\triangleq T_s + O(\mu) = \begin{bmatrix} A_s^T & -\left(C_s^T C_s - \frac{1}{\gamma^2} G_s^T R_s G_s \right) \\ -B_s W_s B_s^T & -A_s \end{bmatrix} + O(\mu), \end{aligned} \tag{2.36}$$

$$\begin{aligned} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} &= T_4 + \mu M T_2 = T_4 + O(\mu) \\ &= \begin{bmatrix} A_2^T & -\left(C_2^T C_2 - \frac{1}{\gamma^2} G_2^T R G_2 \right) \\ -B_2 W B_2^T & -A_2 \end{bmatrix}, \end{aligned} \tag{2.37}$$

then it follows by perturbing the coefficients of the AREs (2.34), (2.35), we get the following $\mathcal{H}_\infty$ symmetric filter AREs:

$$A_s \widetilde{P}_s^{(0)} + \widetilde{P}_s^{(0)} A_s^T - \widetilde{P}_s^{(0)} \left(C_s^T C_s - \frac{1}{\gamma^2} G_s^T R_s G_s \right) \widetilde{P}_s^{(0)} + B_s W_s B_s^T = 0, \tag{2.38}$$

$$A_2 \widetilde{P}_f^{(0)} + \widetilde{P}_f^{(0)} A_2^T - \widetilde{P}_f^{(0)} \left(C_2^T C_2 - \frac{1}{\gamma^2} G_2^T R G_2 \right) \widetilde{P}_f^{(0)} + B_2 W B_2^T = 0. \tag{2.39}$$

Assumption 2.2.1 is sufficient to guarantee the existence of a numerically convergent iterative solution to the ARE (2.39). Similarly, the following assumption is sufficient to guarantee the existence of a numerically convergent positive-definite stabilizing solution of the ARE (2.38).

Assumption 2.2.2. *The triple* $(A_s, C_s, \sqrt{B_s W_s B_s^T})$ *is controllable and observable.*

Next, we consider the decomposition of the filter (2.28). We apply again the Chang decoupling transformation

$$T_F = \begin{bmatrix} I - \mu H L & -\mu H \\ L & I \end{bmatrix}, \quad T_F^{-1} = \begin{bmatrix} I & \mu H \\ -L & I - \mu H L \end{bmatrix},$$

for some mtrices H and L on the closed-loop filter matrix

$$\begin{bmatrix} (A_1 - K_1 C_1) & (A_{12} - K_1 C_2) \\ \frac{1}{\mu}(A_{21} - K_2 C_1) & \frac{1}{\mu}(A_2 - K_2 C2) \end{bmatrix}$$

to obtain the decoupling equations

$$(A_2 - K_2C_2)L - (A_{21} - K_2 - K_2C_1) - \mu[(A_1 - K_1C_1) - (A_{12} - K_1C_2)L] = 0 \tag{2.40}$$

$$-H(A_2 - K_2C_2) + (A_{12} - K_1C_2) - \mu HL(A_{12} - K_1C_2) + \mu[(A_1 - K_1C_1) - (A_{12} - K_1C_2)L]H = 0. \tag{2.41}$$

The unique solution of the above algebraic equations exists under the assumption that the matrix $(A_2 - K_2C_2)$ is nonsingular. This solution can also be obtained by using Newton's method starting with the following initial conditions:

$$\begin{aligned} L^{(0)} &= (A_2 - K_2C_2)^{-1}(A_{21} - K_2C_1), \\ M^{(0)} &= (A_{12} - K_1C_2)(A_2 - K_2C_2)^{-1}. \end{aligned}$$

Thus, application of T_F^{-1} to (2.28) results in the following decomposed filter equations

$$\mathbf{F}^l_{6,\mu}: \begin{cases} \begin{bmatrix} \dot{\hat{\zeta}}_s \\ \dot{\hat{\zeta}}_f \end{bmatrix} & = T_F^{-1} \begin{bmatrix} (A_1 - K_1C_1) & (A_{12} - K_1C_2) \\ \frac{1}{\mu}(A_{21} - K_2C_1) & \frac{1}{\mu}(A_2 - K_2C2) \end{bmatrix} T_F \begin{bmatrix} \hat{\zeta}_s \\ \hat{\zeta}_f \end{bmatrix} + \\ & \quad T_F^{-1} \begin{bmatrix} K_1 \\ \frac{1}{\mu}K_2 \end{bmatrix} y \\ & \triangleq \begin{bmatrix} a_s & 0 \\ 0 & \frac{1}{\mu}a_f \end{bmatrix} \begin{bmatrix} \hat{\zeta}_s \\ \hat{\zeta}_f \end{bmatrix} + \begin{bmatrix} K_s \\ \frac{1}{\mu}K_f \end{bmatrix} y. \end{cases}$$

The filter coefficients and gain matrices are also related to the aggregate ones (2.26) by

$$\begin{aligned} a_s &= (A_1 - K_1C_1) - (A_{12} - K_1C_2)L \\ a_f &= (A_2 - K_2C_1) + \mu L(A_{12} - K_1C_2) \\ K_s &= K_1 - HK_2 - \mu HLK_1 \\ K_f &= K_2 + \mu LK_1. \end{aligned}$$

The above represent the independent pure-slow and pure-fast filters. Due to complete independence, the slow and fast signals can be processed with different sampling rates in contrast with the original full-order filter (2.19), (2.28) which requires the fast sampling rate for processing of both.

2.3 Review of $\mathcal{H}_2$ Filtering for Linear Descriptor Systems

In this section, we review corresponding Kalman or $\mathcal{H}_2$ filtering results for linear descriptor or singular systems. The results presented here are mainly from (Darouach, 1997). We consider the following LTI singular system

$$\Sigma^l_{des}: \begin{cases} E\dot{x} = Ax + Bu + w, & x(t_0) = x_0 \\ y = Cx + v, \end{cases} \tag{2.42}$$

where $E \in \Re^{r\times n}$, with $rank(E) = q \leq n$, $A \in \Re^{r\times n}$, $B \in \Re^{r\times p}$, $C \in \Re^{m\times n}$, and all the variables have their previous meanings. Moreover, w and v are zero-mean Gaussian white noise processes with joint covariance matrices

$$\mathsf{E}\left\{\begin{bmatrix} w(t) \\ v(t) \end{bmatrix}\begin{bmatrix} w^T(\tau) & v^T(\tau)\end{bmatrix}\right\} = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}\delta(t-\tau). \tag{2.43}$$

In addition, we assume the initial condition x_0 is also a Gaussian random variable with

$$\mathsf{E}\{x_0\} = \bar{x}_0, \quad \mathsf{E}\{x_0 x_0^T\} = P_0.$$

The following assumptions will also be required in the sequel.

Assumption 2.3.1. *We assume the following on the system matrices*

(i)

$$rank\begin{bmatrix} E & A \\ 0 & E \\ 0 & C \end{bmatrix} - rank\ E = n, \tag{2.44}$$

(ii)

$$rank\begin{bmatrix} E & A & Q & S \\ 0 & C & S^T & R \\ 0 & E & 0 & 0 \\ 0 & 0 & E^T & 0 \end{bmatrix} = r + m + 2q, \tag{2.45}$$

(iii) $rank(sE - A) = r$ *for almost all* $s \in \mathbf{C}$.

Assumption 2.3.1 (i) is necessary for the Y-observability of the system (if E is square). Whereas Assumption (ii) generalizes the condition $R > 0$ for a standard system (with $E = I$), and finally Assumption (iii) guarantees the system (2.42) is impulse-free.

Since $rank(E) = q$, there exist two nonsingular matrices U and V of appropriate dimensions such that $Uw = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$, $x = V \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$,

$$UEV = \begin{bmatrix} I_q & 0 \\ 0 & 0 \end{bmatrix}, \quad UAV = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}, \quad UB = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad CV = [C_1 \;\; C_2],$$

and the system (2.42) can be represented in this new coordinates as

$$\Sigma_{des}^l : \begin{cases} \dot{x}_1 &= A_1x_1 + A_2x_2 + B_1u + w_1 \\ 0 &= A_3x_1 + A_4x_2 + B_2u + w_2 \\ y &= C_1x_1 + C_2x_2 + v. \end{cases} \tag{2.46}$$

Next, under Assumption 2.3.1 (ii), there exists (Darouach, 1997) a nonsingular transformation $\Gamma \in \Re^{(r-q+m)\times(r-q+m)}$ such that

$$\Gamma \begin{bmatrix} A_4 \\ C_2 \end{bmatrix} = \begin{bmatrix} \Psi_1 \\ \Psi_2 \end{bmatrix} \tag{2.47}$$

where $\Psi_1 \in \Re^{(n-q)\times(n-q)}$ is nonsingular. Define also the following nonsingular matrix

$$T = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} = \begin{bmatrix} \Psi_1^{-1} - \Omega\Psi_2\Psi_1^{-1} & \Omega \\ -\Psi_2\Psi_1^{-1} & I_{r+p-n} \end{bmatrix} \tag{2.48}$$

where Ω is an arbitrary matrix of appropriate dimension, which must be chosen so that the error covariance matrix of the algebraic model is minimized. Then, premultiplying (2.46)(ii) and (iii) by the nonsingular matrix $T\Gamma$ and substituting x_2 into (2.46)(i), we obtain the following equivalent system

$$\bar{\Sigma}_{des}^l \begin{cases} \dot{x}_1 &= \Phi x_1 + \bar{B}\bar{u} + G_1\omega \\ x_2 &= \bar{H}x_1 + \bar{D}\bar{u} + G_2\omega \\ \bar{y} &= \bar{C}x_1 + G_3\omega, \end{cases} \tag{2.49}$$

where

$$\bar{H} = -T_1\Gamma \begin{bmatrix} A_3 \\ C_1 \end{bmatrix}, \quad \bar{C} = T_2\Gamma \begin{bmatrix} A_3 \\ C_1 \end{bmatrix}, \quad \bar{u} = \begin{bmatrix} u \\ y \end{bmatrix}, \quad \omega = \begin{bmatrix} w_1 \\ w_2 \\ v \end{bmatrix},$$

$$\bar{D} = T_1\Gamma \begin{bmatrix} -B_2 & 0 \\ 0 & I_p \end{bmatrix}, \quad \Phi = A_1 + A_2\bar{H}, \quad G_1 = [I_q \quad -A_2T_1\Gamma],$$
$$G_2 = [0 \quad -T_1\Gamma], \quad \bar{B} = \bar{B}_1 + A_2\bar{D}, \quad G_3 = [0 \ \ T_2\Gamma],$$
$$\bar{y} = T_2\Gamma \begin{bmatrix} -B_2u \\ y \end{bmatrix}, \quad \bar{B}_1 = [B_1 \ \ 0],$$

and

$$\mathsf{E}\{\omega(t)\omega^T(\tau)\} = \begin{bmatrix} U & 0 \\ 0 & I_p \end{bmatrix} \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} U^T & 0 \\ 0 & I_p \end{bmatrix} \delta(t-\tau) = \Pi\delta(t-\tau). \tag{2.50}$$

Equation (2.50) indicates that the noise terms are correlated. For simplicity, we would find an equivalent representation in which they are uncorrelated using a sequence of transformations. Accordingly, define

$$\Lambda = \mathsf{E}\{\omega(t)(G_3\omega(\tau))^T\} = \Pi G_3^T\delta(t-\tau), \quad \bar{R} = G_3\Pi G_3^T.$$

Then the following result can be proven (Darouach, 1997)

Lemma 2.3.1. *Under Assumption 2.3.1 (i), (iii), the matrix $\bar{R}$ is positive definite.*

Further, define

$$\begin{aligned} \bar{A} &= \Phi - G_1\Lambda\bar{R}^{-1}\bar{C}, \\ \eta &= (G_1 - G_1\Lambda\bar{R}^{-1}G_3)\omega. \end{aligned}$$

Then

$$\mathsf{E}\{\eta(t)\eta(\tau)^T\} = (G_1 - G_1\Lambda\bar{R}^{-1}G_3)\Pi((G_1 - G_1\Lambda\bar{R}^{-1}G_3)^T\delta(t-\tau) = \bar{Q}\delta(t-\tau), \tag{2.51}$$

and the system (2.46) becomes equivalently

$$\widetilde{\Sigma}^l_{des} \begin{cases} \dot{x}_1 &= \bar{A}x_1 + \bar{B}\bar{u} + G_1\Lambda\bar{R}^{-1}\bar{y} + \eta \\ \bar{y} &= \bar{C}x_1 + G_3\omega, \end{cases} \tag{2.52}$$

with η and $G_3\omega$ uncorrelated.

The filter can now be designed based on the classical Kalman filtering theory for uncorrelated noise terms. Accordingly, consider the following configuration

$$\mathbf{F}^l_{des4} : \left\{ \dot{\hat{x}}_1 = \bar{A}\hat{x}_1 + \bar{B}\bar{u} + G_1\Lambda\bar{R}^{-1}\bar{y} + K(\bar{y} - \bar{C}\hat{x}_1), \right. \tag{2.53}$$

where the gain K is given by

$$K = P_1\bar{C}^T\bar{R}^{-1}$$

and P_1 satisfies the Riccati ordinary differential-equation

$$\dot{P}_1(t) = \bar{A}P_1(t) + P_1(t)\bar{A}^T + \bar{Q} - P_1(t)\bar{C}^T\bar{R}^{-1}\bar{C}P_1(t). \tag{2.54}$$

Then, an unbiased estimate for x_2 is determined from (2.49) as

$$\hat{x}_2 = \mathsf{E}\{x_2\} = \bar{H}\hat{x}_1 + \bar{D}\bar{u} \tag{2.55}$$

with error covariance

$$P_2 = \bar{H}P_1\bar{H}^T + G_2\Pi G_2^T. \tag{2.56}$$

This estimate is optimal if the "trace" of the error covariance P_2 above is minimal. Further, it can be shown that (Darouach, 2008), the result above is independent of Ω and hence the transformation T_1.

In addition, the steady-state filter gain $\bar{K}$ is obtained as the limiting value of the gain K which inturn is determined by the limiting value $\bar{P}_1 = \lim_{t\to\infty} P_1(t)$. This limit if it exists, must satisfy the ARE

$$\bar{A}\bar{P}_1 + \bar{P}_1\bar{A}^T + \bar{Q} - \bar{P}_1\bar{C}^T\bar{R}^{-1}\bar{C}\bar{P}_1 = 0. \tag{2.57}$$

It can then be shown that, under the conditions of Assumption 2.3.1 and some additional structural assumptions, the above ARE has a positive definite stabilizing solution. The following results have been established (Darouach, 2008).

Lemma 2.3.2. *The pair $(\bar{A}, \bar{C})$ is detectable if and only if*

$$rank \begin{bmatrix} sE - A \\ C \end{bmatrix} = n, \ \forall s \in \mathbf{C}, \quad Real(s) \geq 0. \tag{2.58}$$

The following theorem gives necessary and sufficient conditions for the convergence and stability of the filter. They are based on the concept of strong and stabilizing solutions of the ARE (2.57). Briefly, a positive definite solution of the ARE is a strong solution if it is such that the transition matrix of $(\bar{A} - \bar{P}_1\bar{C}^T\bar{R}^{-1}\bar{C})$ has all its eigenvalues λ satisfying $\Re(\lambda) \leq 0$, and it is a stabilizing solution if $(\bar{A} - \bar{P}_1\bar{C}^T\bar{R}^{-1}\bar{C})$ has all its eigenvalues σ satisfying $Re(\sigma) < 0$.

Theorem 2.3.1. *Under Assumption 2.3.1, the following hold:*

(i) The ARE (2.57) has a unique strong solution if and only if (2.58) is satisfied.

(ii) The strong solution is the only nonnegative definite solution of the ARE if and only if (2.58) is satisfied, and the pencil $\begin{bmatrix} A - \lambda E & \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}^{1/2} \\ C & \end{bmatrix}$ *has full-row rank for all finite complex λ satisfying $\Re(\lambda) \geq 0$.*

(iii) The strong solution of the ARE is stabilizing if and only if (2.58) is satisfied, and the pencil $\begin{bmatrix} A - \lambda E & \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}^{1/2} \\ C & \end{bmatrix}$ *has full-row rank for all finite complex λ satisfying*

$\Re(\lambda) = 0$. *If in addition this rank condition is satisfied, for* $\Re(\lambda) < 0$, *then the strong solution is also positive definite.*

Theorem 2.3.2. *Suppose:*

(i) $P_1(0) - \bar{P}_1 \geq 0$, *then* $\lim_{t\to\infty} P_1(t) = \bar{P}_1$, *if and only if (2.58) is satisfied;*

(ii) $P_1(0) > 0$, *then* $\lim_{t\to\infty} P_1(t) = \bar{P}_1$ *exponentially fast, if and only if, (2.58) is satisfied and the pencil* $\begin{bmatrix} A - \lambda E & \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}^{1/2} \\ C & \end{bmatrix}$ *has full-row rank for all finite complex* λ *satisfying* $Real(\lambda) = 0$.

In the next section, we discuss the $\mathcal{H}_\infty$ problem.

2.4 Review of $\mathcal{H}_\infty$ Filtering for Linear Descriptor Systems

In this section, we review some results on the $\mathcal{H}_\infty$ filtering probblem for linear descriptor systems (Xu, 2003b). We consider the following descriptor system

$$\Sigma^l_{des} : \begin{cases} E\dot{x} &= Ax + Bw, \quad x(t_0) = x_0 \\ y &= Cx + Dw \\ z &= Lx, \end{cases} \tag{2.59}$$

where $x \in \Re^n$ is the system state vector, $y \in \Re^m$ is the measured output, $z \in \Re^s$ is the controlled output or penalty variable, $w \in \mathcal{L}_2([t_0, \infty), \Re^r)$ is the noise/disturbance vector, $E \in \Re^{n\times n}$ and $rank(E) = q \leq n$ is the singular matrix of the system, while A, B, C, D and L are real constant matrices of appropriate dimensions. To proceed, we adopt the following definition.

Assumption 2.4.1. *The disturbance-free system* $E\dot{x} = Ax$ *is admissible, i.e., the following hold:*

1. *the system is regular, i.e.,* $det(sE - A) \not\equiv 0$ *identically;*

2. *the system is impulse-free, i.e., $deg(det(sE-A)) = rank(E)$ $\forall s \in \mathbf{C}$;*

3. *the system is stable or (E, A) is Hurwitz., i.e., the roots of $det(sE-A) = 0$ have negative real parts.*

The following Lemmas will be required in the sequel.

Lemma 2.4.1. *(Xu, 2003b) Consider the system (2.59) and let the transfer function from w to z be $G_{zw}(s) = L(sE-A)^{-1}B$. Then, the following statements are equivalent;*

(S1) the system with $w \equiv 0$ is sdmissible and $\|G(s)\|_\infty < \gamma$;

(S2) there exists a matrix P satisfying the following LMIs:

$$E^T P = P^T E \geq 0 \tag{2.60}$$

$$\begin{bmatrix} A^T P + P^T A & P^T B & L^T \\ B^T P & -\gamma^2 I & 0 \\ L & 0 & -I \end{bmatrix} < 0. \tag{2.61}$$

Lemma 2.4.2. *(Boyd, 1994) (Schur complements for nonstrict LMI). The matrix inequality*

$$\begin{bmatrix} Z_1 & Z_2 \\ Z_2^T & Z_3 \end{bmatrix} \geq 0$$

holds if and only if

$$Z_3 \geq 0, \quad Z_1 - Z_2 Z_3^+ Z_2^T \geq 0, \quad Z_2(I - Z_3 Z_3^+) = 0.$$

We consider the following filter for the system

$$\Sigma_{desf}^l : \begin{cases} E\dot{\hat{x}} &= A_f \hat{x} + B_f y \\ \hat{z} &= C_f \hat{x} \end{cases} \tag{2.62}$$

where $\hat{x} \in \Re^{\bar{n}}$ is the filter state, $\hat{z} \in \Re^{\bar{s}}$ is the estimated output of the filter, while A_f, B_f, C_f, are real constant filter matrices of appropriate dimensions which are to be determined.

Letting

$$e = [x^T \quad \hat{x}^T]^T, \quad \tilde{z} = z - \hat{z},$$

the combined closed-loop system (2.59), (2.62) can be represented as

$$\tilde{\Sigma}^l_{desc} : \begin{cases} E_c \dot{e} &= A_c e + B_c w \\ \tilde{z} &= L_c e \end{cases} \tag{2.63}$$

where

$$E_c = \begin{bmatrix} E & 0 \\ 0 & E_f \end{bmatrix}, \quad A_c = \begin{bmatrix} A & 0 \\ B_f C & A_f \end{bmatrix}, \quad B_c = \begin{bmatrix} B \\ B_f D \end{bmatrix}, \quad L_c = [L \quad -C_f].$$

The problem can then be stated as follows. For a given $\gamma > 0$, find the filter matrices A_f, B_f and C_f such that the system (2.63) is admissible and the $\mathcal{H}_\infty$ norm of the system transfer function from w to $\tilde{z}$, denoted as $T_{\tilde{z}w}$, satisfies the constraint $\|T_{\tilde{z}w}\|_\infty < \gamma$ for all $w \in \mathcal{L}_2[t_0, \infty)$.

The following result then gives a solution to the problem.

Theorem 2.4.1. *Consider the systems (2.59) and suppose that it satisfies Assumption 2.4.1. Then there exists a filter of the form Σ^l_{desf} that solves the $\mathcal{H}_\infty$ filtering problem for the system if and only if, there exists matrices X, Y, Φ, Ψ and Υ such that the following LMIs are satisfied*

$$E^T X = X^T E \geq 0 \tag{2.64}$$

$$E^T Y = Y^T E \geq 0 \tag{2.65}$$

$$E^T (X - Y) \geq 0 \tag{2.66}$$

$$\begin{bmatrix} A^TY+Y^TA & A^TX+Y^TA+C^T\Psi^T+\Phi^T & Y^TB & L^T-\Upsilon^T \\ X^TA+A^TY+\Psi C+\Phi & X^TA+A^TX+\Psi C+C^T\Phi^T & X^TB+\Psi D & L^T \\ B^TY & B^TX+D^T\Psi^T & -\gamma^2 I & 0 \\ L-\Upsilon & L & 0 & -I \end{bmatrix} < 0. \quad (2.67)$$

Moreover, in this case, there exist nonsingular matrices S, $\tilde{S}$, W, and $\tilde{W}$ such that

$$E^T\hat{S} = S^TE \qquad (2.68)$$

$$EW = \tilde{W}E^T \qquad (2.69)$$

$$XY^{-1} = I - \tilde{S}W \qquad (2.70)$$

$$Y^{-1}X = I - \tilde{W}S, \qquad (2.71)$$

and the filter matrices are given by

$$E_f = E, \quad A_f = S^{-T}\Phi Y^{-1}W^{-1}, \quad B_f = S^{-T}\Psi, \quad C_f = \Upsilon Y^{-1}W^{-1}$$

Proof: We only give the proof of sufficiency of the Theorem. The necessity part can be found in (Xu, 2003b). (Sufficiency:) Suppose (2.64)-(2.67) hold. Then, we show that there always exist nonsingular matrices S, $\tilde{S}$, W, and $\tilde{W}$ such that (2.68)-(2.71) hold. Accordingly, we first show that the matrix Y satisfying (2.65)-(2.67) is nonsingular. Otherwise, $\exists\eta \neq 0$ such that $Y\eta = 0$, and therefore $\eta^T(A^TY + Y^TA)\eta = 0$. But this contradicts the fact that (2.67) implies $A^TY + Y^TA < 0$. Furthermore, we may assume also without loss of generality that $Y - X$ is nonsingular. Otherwise, we can choose $\hat{Y} = (1-\alpha)Y$, $\alpha > 0$ a sufficiently small number that is not an eigenvalue of $I - XY^{-1}$ and such that $\hat{Y}$ is nonsingular and satisfies (2.67). Then it follows that (2.65), (2.66) are also satisfied by this $\hat{Y}$ and $\hat{Y} - X$ is nonsingular. Thus, we conclude that, we can aways find a nonsingular Y that satisfies (2.65)-(2.67). Moreover, this also implies that $I - XY^{-1}$ and $I - Y^{-1}X$ are also nonsingular.

Next, choose two nonsingular matrices M and N such that

$$E = M \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

Then, X and Y in (2.64, (2.65) can be written as

$$X = M^{-T} \begin{bmatrix} X_1 & 0 \\ X_2 & X_3 \end{bmatrix} N, \quad Y = M^{-T} \begin{bmatrix} Y_1 & 0 \\ Y_2 & Y_3 \end{bmatrix} N,$$

where $X_1 = X_1^T \geq 0$, $Y_1 = Y_1^T > 0$. Moreover,

$$Y^{-1} = N^{-1} \begin{bmatrix} \hat{Y}_1 & 0 \\ \hat{Y}_2 & \hat{Y}_3 \end{bmatrix} M^T,$$

where $\hat{Y}_1 = \hat{Y}_1^T = Y_1^{-1} > 0$. Now set

$$S = M^{-T} \begin{bmatrix} S_1 & 0 \\ S_2 & S_3 \end{bmatrix} N, \quad \tilde{S} = M^{-T} \begin{bmatrix} \tilde{S}_1 & 0 \\ \tilde{S}_2 & \tilde{S}_3 \end{bmatrix} N, \tag{2.72}$$

$$W = N^{-1} \begin{bmatrix} W_1 & 0 \\ W_2 & W_3 \end{bmatrix} M^T, \quad \tilde{W} = N^{-1} \begin{bmatrix} \tilde{W}_1 & 0 \\ \tilde{W}_2 & \tilde{W}_3 \end{bmatrix} M^T, \tag{2.73}$$

and where the matrices S_i, $\tilde{S}_i$, W_i, $\tilde{W}_i$, $i = 1, 2, 3$ are selected to satisfy

$$S_1^T = \tilde{S}_1, \quad W_1 = \tilde{W}_1^T, \tag{2.74}$$

$$\begin{bmatrix} \tilde{S}_1 W_1 & 0 \\ \tilde{S}_2 W_1 + \tilde{S}_3 W_2 & \tilde{S}_3 W_3 \end{bmatrix} = \begin{bmatrix} I - X_1 \hat{Y}_1 & 0 \\ -X_2 \hat{Y}_1 - X_3 \hat{Y}_2 & I - X_3 \hat{Y}_3 \end{bmatrix}, \tag{2.75}$$

$$\begin{bmatrix} W_1 \tilde{S}_1 & 0 \\ \tilde{W}_2 S_1 + \tilde{W}_3 S_2 & \hat{W}_3 S_3 \end{bmatrix} = \begin{bmatrix} I - \hat{Y}_1 X_1 & 0 \\ -\hat{Y}_2 X_1 - \hat{Y}_3 X_3 & I - \hat{Y}_3 X_3 \end{bmatrix}. \tag{2.76}$$

Using these equations, it can be verified that S, $\tilde{S}$, W and $\tilde{W}$ given by (2.72), (2.73) satisfy ((2.68)-(2.71). Moreover, the nonsingularity of $I - XY^{-1}$ and $I - \hat{Y}^{-1}X$ implies that the matrices S, $\tilde{S}$, W and $\tilde{W}$ are nonsingular too.

Next, we show that the error systems $\tilde{\Sigma}^l_{desc}$ is admissible and the filter Σ^l_{desf} is also admissible with $\|T_{\tilde{z}w}\|_\infty < \gamma$ for all $w \in \mathcal{L}_2[t_0, \infty)$.

Define

$$\Pi_1 = \begin{bmatrix} \bar{Y} & I \\ W & 0 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} I & X \\ 0 & S \end{bmatrix},$$

where $\bar{Y} = Y^{-1}$. Then, clearly, both Π_1 and Π_2 are nonsingular. Also, setting

$$\hat{P} = \Pi_2 \Pi_1^{-1} \triangleq \begin{bmatrix} X & \tilde{S} \\ S & -\Gamma \end{bmatrix},$$

where $\Gamma = S\bar{Y}W^{-1}$, we see that $\hat{P}$ is nonsingular as well. Moreover, from (2.68)-(2.71), we have

$$\begin{aligned} E^T\Gamma &= E^T S\bar{Y}W^{-1} = W^{-T}(W^T E^T S\bar{Y})W^{-1} = W^{-T}E\tilde{W}S\bar{Y}W^{-1} \\ &= W^{-T}E(I - \bar{Y}X)\bar{Y}W^{-1} = W^{-T}E(\bar{Y} - \bar{Y}X\bar{Y})W^{-1} \\ &= W^{-T}E\bar{Y}(Y - X)\bar{Y}W^{-1} = W^{-T}\bar{Y}^T[E^T(Y - X)]\bar{Y}W^{-1}. \end{aligned}$$

Next, by (2.64)-(2.66), we have

$$E^T\Gamma = \Gamma^T E \le 0, \tag{2.77}$$

and therefore

$$\hat{E}^T\hat{P} = \hat{P}\hat{E}, \tag{2.78}$$

where $\hat{E} = diag\{E, E\}$. Now, noting that Γ is nonsingular, and using (2.77), (2.68), (2.70), we have

$$\begin{aligned} E^T X + E^T\tilde{S}\Gamma^{-1}(\Gamma^{-T}E^T)^+\Gamma^{-T}E^T S &= E^T X + S^T E\Gamma^{-1}(\Gamma^{-T}E^T)^+\Gamma^{-T}E^T S \\ &= E^T X + S^T\Gamma^{-T}E^T(\Gamma^{-T}E^T)^+\Gamma^{-T}E^T S \\ &= E^T X + S^T\Gamma^{-T}(\Gamma^{-T}E^T)^+\Gamma^{-T}E^T S \\ &= E^T X + S^T\Gamma^{-T}E^T S = E^T X + S^T E\Gamma^{-1}S \\ &= E^T X + E^T\tilde{S}\Gamma^{-1}S = E^T(X + \tilde{S}\Gamma^{-1}S) \end{aligned}$$

$$
\begin{aligned}
&= E^T(X + \tilde{S}WY) = E^T[X + (I - X\bar{Y})Y] \\
&= E^T Y \geq 0. \qquad (2.79)
\end{aligned}
$$

Furthermore, since $E\Gamma^{-1}$ is symmetric, we obtain

$$
\begin{aligned}
E^T\tilde{S}\Gamma^{-1}[I - (-\Gamma^{-T}E^T)(-\Gamma^{-T}E^T)^+] &= S^T E\Gamma^{-1}[I - (E\Gamma^{-1})^T((E\Gamma^{-1})^+)^T] \\
&= S^T E\Gamma^{-1}[I - \left((E\Gamma^{-1})^+(E\Gamma^{-1})\right)^T] \\
&= S^T E\Gamma^{-1}[I - (E\Gamma^{-1})^+(E\Gamma^{-1})] \\
&= S^T[E\Gamma^{-1} - (E\Gamma^{-1})(E\Gamma^{-1})^+(E\Gamma^{-1})] \\
&= 0. \qquad (2.80)
\end{aligned}
$$

By (2.79), (2.80) and Lemma 2.4.2, we deduce that

$$
\begin{bmatrix} E^T X & E^T\tilde{S}\Gamma^{-1} \\ \Gamma^{-T}E^T S & -\Gamma^{-T}E^T \end{bmatrix} \geq 0. \qquad (2.81)
$$

Premultiplying (2.81) by $diag\{I, \Gamma^T\}$ and postmultiplying by $diag\{I, \Gamma\}$, gives

$$
\begin{bmatrix} E^T X & E^T\tilde{S} \\ E^T S & -E^T\Gamma \end{bmatrix} \geq 0. \qquad (2.82)
$$

Noting (2.78), (2.82) can be written as

$$
\hat{E}^T\hat{P} = \hat{P}^T\hat{E} \geq 0. \qquad (2.83)
$$

On the other hand, pre-multiplying (2.67) by $diag\{\bar{Y}^T, I, I, I\}$ and post-multiplying by $diag\{\bar{Y}, I, I, I\}$, we have

$$
\left[\begin{array}{ccc}
\bar{Y}^T A^T + A\bar{Y} & \bar{Y}^T A^T X + A + \bar{Y}^T C^T \Psi^T + \bar{Y}^T \Phi^T & B \\
X^T A\bar{Y} + A^T + \Psi C\bar{Y} + \Phi\bar{Y} & X^T A + A^T X + \Psi C + C^T \Psi^T & X^T B + \Psi D \\
B^T & B^T X + D^T \Psi^T & -\gamma^2 I \\
L\bar{Y} - \Upsilon\bar{Y} & L & 0
\end{array} \right.
$$

$$\begin{matrix} \bar{Y}^T L^T - \bar{Y}^T \Upsilon^T \\ L^T \\ 0 \\ -I \end{matrix} \Bigg] < 0. \tag{2.84}$$

The above inequality can also be rewritten as

$$\begin{bmatrix} \Pi_1^T A_c^T \hat{P} \Pi_1 + \Pi_1^T \hat{P}^T A_c \Pi_1 & \Pi_1^T \hat{P}^T B_c & \Pi_1^T L_c^T \\ B_c^T \hat{P} \Pi_1 & -\gamma^2 I & 0 \\ L_c \Pi_1 & 0 & -I \end{bmatrix} < 0. \tag{2.85}$$

Again, pre-multiplying (2.85) by $diag\{\Pi^{-T}, I, I\}$ and post-multiplying by $diag\{\Pi^{-1}, I, I\}$, we obtain

$$\begin{bmatrix} A^T c\hat{P} + \hat{P}^T A_c & \hat{P}^T B_c & L_c^T \\ B_c^T \hat{P} & -\gamma^2 I & 0 \\ L_c & 0 & -I \end{bmatrix} < 0. \tag{2.86}$$

Combining (2.86), (2.83) and using Lemma 2.4.1 the result follow. □

2.5 Motivations and Research Objectives

As mentioned in the Introduction and discussed in the previous sections, the filtering problem for linear singularly-perturbed systems has been considered by many authors, and various types of filters have been proposed; including, composite, decomposition and reduced-order filters. However, to the best of our knowledge, the problem for affine nonlinear singularly-perturbed systems has not been considered by any authors. Although, the dynamic output-feedback problem for a class of systems has been considered by a handful of authors (Assawinchaichote, 2004b), and the same authors have also considered the filtering problem for a class of stochastic Tagaki-sugeno fuzzy nonlinear systems (Assawinchaichote, 2004a). Therefore, in this section, we outline as part of our research objectives to discuss the above problem for both continuous-time and discrete-time nonlinear singularly-perturbed systems.

Similarly, as mentioned in the Introduction and reviewed in the previous sections, various authors have also considered the observer design problem for linear descriptor systems (Dai 1989), (Dai 1988), (Darouach, 1995)-(Darouach, 2008), (El-Tohami, 1987)-(Fahmy, 1989), (Hou, 1995), (Koenig, 1995), (Minamide, 1989), (Paraskevopoulos, 1992), (Sun, 2007), (Uetake, 1989), (Zhou, 2008). Kalman-Luenberger type full-order and reduced-order observers have extensively been studied, and necessary and sufficient conditions for the solvability of the problem have been presented. On the other hand, only recently has there been some attention on the design of observers and filters for nonlinear descriptor systems (Darouach, 2008). In addition, to the observer design problem for linear systems studied in the references above, the Kalman filtering problem has also been discussed (Dai, 1989), (Nikoukhah, 1999), (Nikoukhah, 1992), (Zhou, 2008). Similarly, the output observation design problem for nonlinear systems has also been considered in (Zimmer, 1997), but to the best of our knowledge, the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ filtering problems for more general affine nonlinear descriptor systems has not been discussed in any reference. Therefore, we include this problem also as part of our research objectives.

Notwithstanding the above motivations that we have mentioned, the following are even stronger reasons why we undertake this investigation:

1. Statistical nonlinear filtering techniques developed using minimum-variance, maximum likehood, bayesian methods lead to infinite-dimensional filters that require the solution of evolution equations such as the *Stratonovich equation*, the *Kushner equation* and the *Wong-and-Zakai equation* that are known to have no explicit analytical solution and neither any tractable numerical solutions;

2. The extended-Kalman filter (EKF) is known to be inaccurate for highly nonlinear systems and is difficult to tune and implement;

3. Unscented Kalman filters (UKF) and Particle filters (PFs) are still computationally cumbersome;

4. The methods we develop are simple and utilize the full nonlinear system dynamics as

opposed to local linearization. They rely on finding smooth solutions to *Hamilton-Jacobi equations (HJEs)* that can be found using polynomial approximation or other methods;

5. The methods we develop are the first successful application of *Hamilton-Jacobi theory* to nonlinear filtering - earlier not too successful methods include the following

   (i) Mortenson R. E. (Mortenson, 1968): "Maximum likelihood recucursive nonlinear filtering," leads to a highly complicated HJE that involves a rank 3 tensor;

   (ii) Berman N. & Shaked U. (Berman, 1996) and Shaked & Berman (Shaked, 1995): "$\mathcal{H}_\infty$ nonlinear filtering," lead to a filter in which the gain matrix depends on the original state and hence is practically not implementable except for linear systems;

   (iii) Nguang S. K. & Fu M. (Nguang, 1996): "Robust $\mathcal{H}_\infty$ nonlinear filtering," also leads to filter whose gain matrix depends on the original state and hence is practically not implementable except for linear systems;

Therefore, we itemize our research objectives as follows:

1. To solve the $\mathcal{H}_2$ continuous-time and discrete-time filtering problem for affine nonlinear singularly perturbed systems;

2. To solve the $\mathcal{H}_2$ continuous-time and discrete-time filtering problem for affine nonlinear descriptor systems;

3. To solve the $\mathcal{H}_\infty$ continuous-time and discrete-time filtering problem for affine nonlinear singularly-perturbed systems;

4. To solve the $\mathcal{H}_\infty$ continuous-time and discrete-time filtering problem for affine nonlinear descriptor systems.

2.6 Conclusion

In this chapter, we have reviewed the literature on Kalman and $\mathcal{H}_\infty$ filtering for linear singular systems; both singularly-perturbed and descriptor systems. However, only continuous-time results have been discussed, while the discrete-time results can be found in the references cited. We have also outlined our motivation and research objectives for the Dissertation. Moreover, in the subsequent chapters, we shall present results on our initial attempts to the solution to the problems outlined in the research objectives.

CHAPTER 3

$\mathcal{H}_2$ FILTERING FOR SINGULARLY-PERTURBED NONLINEAR SYSTEMS

In this chapter, we discuss the $\mathcal{H}_2$ or Kalman filtering problem for affine singularly perturbed nolinear systems. The extended Kalman-filter (EKF) (or nonlinear $\mathcal{H}_2$-filter) is by far the most widely used tool in this area because of its simplicity and near optimal performance. However, it still suffers from the problem of local linearization and is derived from the basic assumption that the measurement noise signal is white Gaussian. This is however seldom the case. Thus, in this chapter, we present alternative approaches to the EKF in which we consider the full system dynamics. Moreover, $\mathcal{H}_2$ techniques are useful when the system and measurement noise are known to be approximately Gaussian distributed.

Two types of filters will be discussed, and sufficient conditions for the solvability of the problem in terms of Hamilton-Jacobi-Bellman equations (HJBEs) will be presented. Both the continuous-time and the discrete-time problems will be considered. The chapter is organized as follows. In section 2, we discuss the continuous-time problem while in section 3, we discuss the discrete-time problem. Finally, in Section 4, a brief conclusion is given. Moreover, in each section we also give problem definition and other preliminaries. Then, the solution to the problem using decomposition filters and then using aggregate filters are presented in subsequent subsections respectively. Examples are then presented to demonstrate the approach.

3.1 $\mathcal{H}_2$ Filtering for Continuous-Time Systems

In this section, we present preliminary results on the $\mathcal{H}_2$ filtering problem for continuous-time affine nonlinear systems, while in the next section we present the discrete-time results. We begin with the problem definition and other preliminary definitions.

3.1.1 Problem Definition and Preliminaries

The general set-up for studying $\mathcal{H}_2$ filtering problems is shown in Fig. 3.1, where $\mathbf{P}$ is the plant, while $\mathbf{F}$ is the filter. The noise signal $w \in \mathcal{S}$ is in general a bounded spectral-signal (e.g. a Gaussian white-noise signal) which belongs to the set $\mathcal{S}$ of bounded spectral-signals, while $\tilde{z} \in \mathcal{P}$, is a bounded power signal or $\mathcal{L}_2$ signal, which belongs to the space of bounded power-signals. Thus, the induced norm from w to $\tilde{z}$ (the penalty variable to be defined later) is the $\mathcal{L}_2$-norm of the interconnected system $\mathbf{F} \circ \mathbf{P}$, i.e.,

$$\|\mathbf{F} \circ \mathbf{P}\|_{\mathcal{L}_2} \stackrel{\Delta}{=} \sup_{0 \neq w \in \mathcal{S}} \frac{\|\tilde{z}\|_{\mathcal{P}}}{\|w\|_{\mathcal{S}}}, \tag{3.1}$$

and is defined as the $\mathcal{H}_2$-norm of the system for stable plant-filter pair $\mathbf{F} \circ \mathbf{P}$, where

$$\begin{aligned} \mathcal{P} &\stackrel{\Delta}{=} \{w(t) : \; w \in \mathcal{L}_\infty, \; R_{ww}(\tau), \; S_{ww}(j\omega) \text{ exist for all } \tau \text{ and all } \omega \text{ resp.}, \; \|w\|_{\mathcal{P}} < \infty\}, \\ \mathcal{S} &\stackrel{\Delta}{=} \{w(t) : \; w \in \mathcal{L}_\infty, \; R_{ww}(\tau), \; S_{ww}(j\omega) \text{ exist for all } \tau \text{ and all } \omega \text{ resp.}, \; \|S_{ww}(j\omega)\|_\infty < \infty\}, \end{aligned}$$

$$\|z\|_{\mathcal{P}}^2 \stackrel{\Delta}{=} \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \|z(t)\|^2 dt,$$

$$\|w\|_{\mathcal{S}}^2 = \|S_{ww}(j\omega)\|_\infty,$$

and $R_{ww}(\tau)$, $S_{ww}(j\omega)$ are the autocorrelation and power-spectral density matrices of $w(t)$ respectively. Notice also that, $\|(.)\|_{\mathcal{P}}$ and $\|(.)\|_{\mathcal{S}}$ are seminorms.

At the outset, we consider the following affine nonlinear causal state-space model of the plant which is defined on a manifold $\mathcal{X} \subseteq \Re^{n_1+n_2}$ with zero control input:

$$\mathbf{P}_{sp}^a : \begin{cases} \dot{x}_1 &= f_1(x_1, x_2) + g_{11}(x_1, x_2)w; \quad x_1(t_0, \varepsilon) = x_{10} \\ \varepsilon\dot{x}_2 &= f_2(x_1, x_2) + g_{21}(x_1, x_2)w; \quad x_2(t_0, \varepsilon) = x_{20} \\ y &= h_{21}(x_1) + h_{22}(x_2) + k_{21}(x_1, x_2)w, \end{cases} \tag{3.2}$$

where $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathcal{X}$ is the state vector with x_1 the slow state which is n_1-dimensional, and x_2 the fast, which is n_2-dimensional; $w \in \mathcal{W} \subseteq \Re^m$ is an unknown disturbance (or

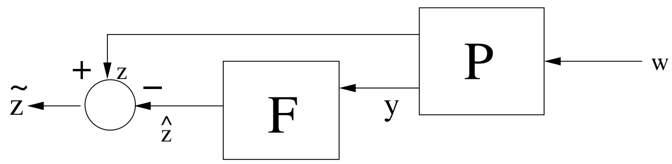

Figure 3.1 Set-up for $\mathcal{H}_2$ Filtering

noise) signal, which belongs to the set $\mathcal{W}$ of admissible exogenous inputs; $y \in \mathcal{Y} \subset \Re^m$ is the measured output (or observation) of the system, and belongs to $\mathcal{Y}$, the set of admissible measured-outputs; while $\varepsilon > 0$ is a small perturbation parameter.

The functions $\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} : \mathcal{X} \to T\mathcal{X} \subseteq \Re^{2(n_1+n_2)}$ [1], $g_{11} : \mathcal{X} \to \mathcal{M}^{n_1 \times m}(\mathcal{X})$, $g_{21} : \mathcal{X} \to \mathcal{M}^{n_2 \times m}(\mathcal{X})$, $h_{21}, h_{22} : \mathcal{X} \to \Re^m$, and $k_{21} : \mathcal{X} \to \mathcal{M}^{m \times m}(\mathcal{X})$, where $\mathcal{M}^{i \times j}$ is the ring of $i \times j$ smooth matrices over $\mathcal{X}$, are real C^∞ functions of x. Furthermore, we assume without any loss of generality that the system (3.2) has an isolated equilibrium-point at $(x_1^T, x_2^T) = (0,0)$ such that $f_1(0,0) = 0$, $f_2(0,0) = 0$, $h_{21}(0,0) = h_{22}(0,0) = 0$. We also assume that there exists a unique solution $x(t, t_0, x_0, w, \varepsilon)$ $\forall t \in \Re$ for the system for all initial conditions $x(t_0) \stackrel{\Delta}{=} x_0 = (x_{10}^T,\ x_{20}^T)^T$, for all $w \in \mathcal{W}$, and all $\varepsilon \in \Re$.

The standard $\mathcal{H}_2$ local filtering/ state estimation problem is defined as follows.

Definition 3.1.1. *(Standard $\mathcal{H}_2$ Local State Estimation (Filtering) Problem). Find a filter,* $\mathbf{F}$*, for estimating the state $x(t)$ or a function of it, $z = h_1(x)$, from observations $\mathsf{Y}_t \stackrel{\Delta}{=} \{y(\tau) : \tau \le t\}$ of $y(\tau)$ up to time t, to obtain the estimate*

$$\hat{x}(t) = \mathbf{F}(\mathsf{Y}_t),$$

such that, the $\mathcal{H}_2$-norm from the input w to some suitable penalty function $\tilde{z}$ is locally minimized for all initial conditions $x_0 \in \mathcal{O} \subset \mathcal{X}$, for all $w \in \mathcal{W} \subset \mathcal{S}$. Moreover, if the filter solves the problem for all $x_0 \in \mathcal{X}$, we say the problem is solved globally.

We shall adopt the following definition of local zero-input observability.

[1] For a manifold M, TM and $T^\star M$ are the tangent and cotangent bundles of M.

Definition 3.1.2. *For the nonlinear system $\mathbf{P}^a_{sp}$, we say that it is locally zero-input observable, if for all states $x_1,\ x_2 \in U \subset \mathcal{X}$ and input $w(.) = 0$,*

$$y(t; x_1, w) \equiv y(t; x_2, w) \Longrightarrow x_1 = x_2,$$

where $y(., x_i, w), i = 1, 2$ is the output of the system with the initial condition $x(t_0) = x_i$. Moreover, the system is said to be zero-input observable if it is locally observable at each $x_0 \in \mathcal{X}$ or $U = \mathcal{X}$.

3.1.2 Decomposition Filters

In this subsection, we present a decomposition approach to the $\mathcal{H}_2$ state estimation problem defined in the previous section, while in the next subsection, we present an aggregate approach. For this purpose, we assume that the noise signal $w \in \mathcal{W} \subset \mathcal{S}$ is a zero-mean Gaussian white-noise process, i.e.,

$$\mathsf{E}\{w(t)\} = 0, \quad \mathsf{E}\{w(t)w^T(\tau)\} = W\delta(t - \tau).$$

Also, the initial conditions $x_1(t_0, \varepsilon) = x_{10},\ x_2(t_0, \varepsilon) = x_{20}$ are assumed to be Gaussian distributed random variables with means given by

$$\mathsf{E}\{x_{10}\} = \bar{x}_{10}, \quad \mathsf{E}\{x_{20}\} = \bar{x}_{20}.$$

We construct two-time scale filters corresponding to the decomposition of the system into a *"fast"* and *"slow"* subsystems. As in the linear case (Chang, 1972), (Gajic, 1994), (Haddad, 1976), we first assume that there exists locally a smooth invertible coordinate transformation (a diffeomorphism)

$$\xi_1 = \varphi_1(x), \quad \varphi_1(0) = 0, \quad \xi_2 = \varphi_2(x), \quad \varphi_2(0) = 0, \ \ \xi_1 \in \Re^{n_1}, \xi_2 \in \Re^{n_2}, \tag{3.3}$$

such that the system (3.2) is locally decomposed into the form

$$
\tilde{\mathbf{P}}^{\mathbf{a}}_{sp} : \begin{cases} \dot{\xi}_1 &= \tilde{f}_1(\xi_1) + \tilde{g}_{11}(\xi)w, \quad \xi_1(t_0) = \varphi_1(x_0) \\ \varepsilon\dot{\xi}_2 &= \tilde{f}_2(\xi_2) + \tilde{g}_{21}(\xi)w; \quad \xi_2(t_0) = \varphi_2(x_0) \\ y &= \tilde{h}_{21}(\xi_1) + \tilde{h}_{22}(\xi_2) + \tilde{k}_{21}(\xi)w. \end{cases} \tag{3.4}
$$

Necessary conditions that such a transformation must satisfy are given by the following proposition.

Proposition 3.1.1. *Consider the nonlinear system (3.2) defined on $\mathcal{X}$. Let (U_1, x), (U_2, ξ), $U_1, U_2 \subset \mathcal{X}$ containing the origin, be two coordinate neighborhoods on $\mathcal{X}$, and consider the problem of finding a local diffeomorphism[2] $\varphi : U_1 \to U_2$, $\xi = \varphi(x)$ so that the system is transformed into the partially decoupled form (3.4) by this coordinate change. Then, the necessary conditions that such a transformation must satisfy are given by the following:*

(i) φ_* *is locally an isomorphism;*

(ii)

$$
\left\langle \frac{\partial}{\partial \xi_j},\ d\left(\varphi^{-1*}\langle f_1 \frac{\partial}{\partial x_1} + \frac{1}{\varepsilon} f_2 \frac{\partial}{\partial x_2},\ d\varphi_i\rangle\right)\right\rangle = 0, \quad i, j = 1, 2,\ i \neq j; \tag{3.5}
$$

(iii)

$$
\left\langle \frac{\partial}{\partial \xi_j},\ \varphi^{-1*} dh_{2i} \right\rangle = 0, \quad i, j = 1, 2, i \neq j, \tag{3.6}
$$

where "$(_*)$", "$(^*)$" are the push-forward and pull-back operators (Boothby, 1975) respectively;

(iv) the following diagrams commute

$$
\begin{tikzcd}
TU_1 \arrow[r, "\varphi_*"] & TU_2 \\
U_1 \arrow[u, "\begin{pmatrix} f_1 \\ \frac{1}{\varepsilon} f_2 \end{pmatrix}"] \arrow[r, "\varphi"] & U_2 \arrow[u, "\begin{pmatrix} \tilde{f}_1 \\ \frac{1}{\varepsilon} \tilde{f}_2 \end{pmatrix}"]
\end{tikzcd}
\qquad
\begin{tikzcd}
\Re^m \arrow[r, "I"] & \Re^m \\
U_1 \arrow[u, "h_{21}"] \arrow[r, "\varphi"] \arrow[d, "h_{22}"] & U_2 \arrow[u, "\tilde{h}_{21}"] \arrow[d, "\tilde{h}_{22}"] \\
\Re^m \arrow[r, "I"] & \Re^m
\end{tikzcd}
$$

[2] *see Ref. (Boothby, 1975) for most of the terminology here.*

Proof: Conditions (i), (ii) and (iii) can be rewritten respectively as

$$
\begin{aligned}
det \begin{bmatrix} \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial x_2} \\ \frac{\partial \varphi_2}{\partial x_1} & \frac{\partial \varphi_2}{\partial x_2} \end{bmatrix} (0) &\neq 0 && (3.7) \\
\frac{\partial}{\partial \xi_2} \left(\frac{\partial \varphi_1}{\partial x_1} f_1(x_1, x_2) + \frac{1}{\varepsilon} \frac{\partial \varphi_1}{\partial x_2} f_2(x_1, x_2) \right) \circ \varphi^{-1}(\xi) &= 0 && (3.8)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial \xi_1} \left(\frac{\partial \varphi_2}{\partial x_1} f_1(x_1, x_2) + \frac{1}{\varepsilon} \frac{\partial \varphi_2}{\partial x_2} f_2(x_1, x_2) \right) \circ \varphi^{-1}(\xi) &= 0 && (3.9) \\
\frac{\partial}{\partial \xi_2} h_{21} \circ \varphi^{-1}(\xi) &= 0 && (3.10) \\
\frac{\partial}{\partial \xi_1} h_{22} \circ \varphi^{-1}(\xi) &= 0. && (3.11)
\end{aligned}
$$

Then, equation (3.7) which corresponds to (i), guarantees that the transformation φ is locally invertible and satisfies the requirements of the Inverse-function Theorem (Sastry, 1999). While equations (3.8), (3.9) and equations (3.10), (3.11) which correspond to conditions (ii), (iii) respectively, guarantee that $\{\tilde{f}_1(\xi_1), \tilde{h}_{21}(\xi_1)\}$, and $\{\tilde{f}_2(\xi_2), \tilde{h}_{22}(\xi_2)\}$ are independent of ξ_2, ξ_1 respectively. Finally, (iv) follows by integrating equations (3.8)-(3.11), and since $\varphi(0) = 0$, $h_{21}(0,0) = 0$, $h_{22}(0,0) = 0$, we get

$$
\begin{aligned}
\left(\frac{\partial \varphi_1}{\partial x_1} f_1(x_1, x_2) + \frac{1}{\varepsilon} \frac{\partial \varphi_1}{\partial x_2} f_2(x_1, x_2) \right) \circ \varphi^{-1}(\xi) &= \tilde{f}_1(\xi_1) && (3.12) \\
\left(\frac{\partial \varphi_2}{\partial x_1} f_1(x_1, x_2) + \frac{1}{\varepsilon} \frac{\partial \varphi_2}{\partial x_2} f_2(x_1, x_2) \right) \circ \varphi^{-1}(\xi) &= \tilde{f}_2(\xi_2) && (3.13) \\
h_{21} \circ \varphi^{-1}(\xi) &= \tilde{h}_{21}(\xi_1) && (3.14) \\
h_{22} \circ \varphi^{-1}(\xi) &= \tilde{h}_{22}(\xi_2). \quad \square && (3.15)
\end{aligned}
$$

We consider an example.

Example 3.1.1. *Consider the following system defined on* $\Re^2 \setminus \{x_1 = 0\}$,

$$
\begin{aligned}
\dot{x}_1 &= -x_1 - \frac{x_2^2}{x_1} + w_0 \\
\varepsilon \dot{x}_2 &= -x_2 + x_2 w_0
\end{aligned}
$$

$$y = x_1 + x_2 + w_0.$$

where w_0 *is a zero-mean Gaussian white noise process. The system has an equilibrium point at* $x = 0$*, but is not defined on the* $x_1 = 0$ *axis. Therefore, it cannot approach* $x = 0$ *along this axis. The transformation*

$$\xi_1 = \varphi_1(x) = \frac{x_1}{x_2}, \quad \xi_2 = \varphi_2(x) = -x_2,$$

on $U_1 = \Re^2 \setminus \{x_2 = 0\}$*, is a local diffeomorphism for the system, and transforms it to*

$$\begin{aligned} \dot{\xi}_1 &= -\xi_1^{-1} + (-\frac{1}{\xi_2} - \xi_1)w_0 \\ \varepsilon\dot{\xi}_2 &= -\xi_2 - \xi_2 w_0 \\ y &= -\xi_1\xi_2 - \xi_2 + w_0. \end{aligned}$$

Similarly, for the set $U_2 = \{x_2 = 0\}$*, we can use the local diffeomorphism*

$$\xi_1 = \tilde{\varphi}_1(x) = x_1, \quad \xi_2 = \tilde{\varphi}_2(x) = x_2,$$

which transforms it to

$$\begin{aligned} \dot{\xi}_1 &= -\xi_1 + w_0 \\ \varepsilon\dot{\xi}_2 &= 0 \\ y &= \xi_1 + w_0. \end{aligned}$$

Remark 3.1.1. *Based on the above example, and since it is too stringent to find a transformation such that* $\tilde{h}_{2j} = \tilde{h}_{2j}(\xi_j), j = 1, 2$*, condition (3.6) equivalently, (3.10), (3.11) can be relaxed or eliminated from the necessary conditions.*

The filter is then designed based on this transformed model with $\mathsf{E}\{w\} = 0$. Accordingly, we propose the following composite filter

$$\mathbf{F}^a_{1c}: \begin{cases} \dot{\hat{\xi}}_1 &= \tilde{f}_1(\hat{\xi}_1) + \hat{L}_1(\hat{\xi}, y)[y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2)]; \quad \hat{\xi}_1(t_0) = \mathsf{E}\{\varphi_1(x_0)\} \\ \varepsilon\dot{\hat{\xi}}_2 &= \tilde{f}_2(\hat{\xi}_2) + \hat{L}_2(\hat{\xi}, y)[y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2)]; \quad \hat{\xi}_2(t_0) = \mathsf{E}\{\varphi_2(x_0)\}. \end{cases} \tag{3.16}$$

where $\hat{\xi} \in \mathcal{X}$ is the filter state, $\hat{L}_1 \in \Re^{n_1 \times m}$, $\hat{L}_2 \in \Re^{n_2 \times m}$ are the filter gains, while all the other variables have their corresponding previous meanings and dimensions. We can then define the penalty variable or estimation error as

$$\tilde{z} = y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2). \tag{3.17}$$

The problem can then be formulated as a dynamic optimization problem with the following cost functional

$$\min_{\substack{\hat{L}_1, \hat{L}_2 \in \Re^{n \times m}, \\ w \in \mathcal{S}, x_0 = 0}} \hat{J}(\hat{L}_1, \hat{L}_2, w) = \mathsf{E}\left\{\frac{1}{2}\int_{t_0}^{\infty} \{\|\tilde{z}\|^2_W\} dt\right\} = \frac{1}{2}\left\{\|\mathbf{F}^a_{1c} \circ \tilde{\mathbf{P}}^{\mathbf{a}}_{sp}\|^2_{\mathcal{H}_2}\right\}_W \tag{3.18}$$

$$s.t. \quad (3.16) \quad \text{and with } w = 0, \ \lim_{t\to\infty}\{\hat{\xi}(t) - \xi(t)\} = 0.$$

To solve the above problem, we form the averaged Hamiltonian function $H : T^\star\mathcal{X} \times T^\star\mathcal{Y} \times \mathcal{W} \times \Re^{n_1 \times m} \times \Re^{n_2 \times m} \to \Re$:

$$\begin{aligned} H(\hat{\xi}, y, w, \hat{L}_1, \hat{L}_2, \hat{V}^T_{\hat{\xi}}, \hat{V}_y) &= \hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)[\tilde{f}_1(\hat{\xi}_1) + \hat{L}_1(\hat{\xi}, y)[y - \tilde{h}_{21}(\hat{\xi}_1) - h_{22}(\hat{\xi}_2))] + \\ &\quad \frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)[\tilde{f}_2(\hat{\xi}_2) + \hat{L}_2(\hat{\xi}, y)[y - \tilde{h}_{21}(\hat{\xi}_1) - h_{22}(\hat{\xi}_2))] + \\ &\quad \hat{V}_y(\hat{\xi}, y)\dot{y} + \frac{1}{2}\|\tilde{z}\|^2_W, \end{aligned} \tag{3.19}$$

for some C^1 function $\hat{V} : \mathcal{X} \times \mathcal{Y} \to \Re$, and where the costate vector $(p_1^T \ \ p_2^T)^T$ is set equal to $(p_1^T \ \ p_2^T)^T = (\hat{V}_{\hat{\xi}} \ \ \hat{V}_y)^T$. Moreover, here and subsequently, $\hat{V}_{\hat{\xi}}$, $\hat{V}_y$ represent row vectors of partial derivatives with respect to $\hat{\xi}$ and y respectively. Completing the squares now for $\hat{L}_1$

and $\hat{L}_2$ in the above expression (3.19), we have

$$\begin{aligned}
H(\hat{\xi}, y, w, \hat{L}_1, \hat{L}_2, \hat{V}_{\hat{\xi}}^T, \hat{V}_y) \;=\;& \hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)\tilde{f}_1(\hat{\xi}_1) + \frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)\tilde{f}_2(\hat{\xi}_2) + \hat{V}_y(\hat{\xi}, y)\dot{y} + \\
& \frac{1}{2}\left\|\hat{L}_1^T(\hat{\xi}, y)\hat{V}_{\hat{\xi}_1}(\hat{\xi}, y) + (y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))\right\|^2 - \\
& \frac{1}{2}\left\|(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))\right\|^2 - \\
& \frac{1}{2}\hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)\hat{L}_1(\hat{\xi}, y)\hat{L}_1^T(\hat{\xi}, y)V_{\hat{\xi}_1}^T(\hat{\xi}, y) + \\
& \frac{1}{2}\left\|\frac{1}{\varepsilon}\hat{L}_2^T(\hat{\xi}, y)\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y) + (y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))\right\|^2 - \\
& \frac{1}{2}\|(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))\|^2 - \\
& \frac{1}{2\varepsilon^2}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)\hat{L}_2(\hat{\xi}, y)\hat{L}_2^T(\hat{\xi}, y)\hat{V}_{\hat{\xi}_2}^T(\hat{\xi}, y) + \frac{1}{2}\|z\|_W^2.
\end{aligned}$$

Thus, setting the optimal gains $\hat{L}_1^\star(\hat{\xi}, y)$, $\hat{L}_2^\star(\hat{\xi}, y)$ as

$$\hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)\hat{L}_1^\star(\hat{\xi}, y) = -(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))^T \tag{3.20}$$

$$\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)\hat{L}_2^\star(\hat{\xi}, y) = -\varepsilon(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))^T \tag{3.21}$$

minimizes the Hamiltonian (3.19). Finally, setting

$$H(\hat{\xi}, y, w, \hat{L}_1^\star, \hat{L}_2^\star, \hat{V}_{\hat{\xi}}^T, \hat{V}_y) = 0,$$

results in the following Hamilton-Jacobi-Bellman equation (HJBE):

$$\begin{aligned}
&\hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)\tilde{f}_1(\hat{\xi}_1) + \tfrac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)\tilde{f}_2(\hat{\xi}_2) + \hat{V}_y(\hat{\xi}, y)\dot{y} - \tfrac{1}{2}\hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)\hat{L}_1(\hat{\xi}, y)\hat{L}_1^T(\hat{\xi}, y)\hat{V}_{\hat{\xi}_1}^T(\hat{\xi}, y) - \\
&\tfrac{1}{2\varepsilon^2}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)\hat{L}_2(\hat{\xi}, y)\hat{L}_2^T(\hat{\xi}, y)\hat{V}_{\hat{\xi}_2}^T(\hat{\xi}, y) + \tfrac{1}{2}(y - \tilde{h}_{21}(\hat{\xi}_1) - \\
&\tilde{h}_{22}(\hat{\xi}_2))^T(W - 2I)(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2)) = 0, \quad \hat{V}(0,0) = 0,
\end{aligned} \tag{3.22}$$

or equivalently the HJBE

$$\begin{aligned}
&\hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)\tilde{f}_1(\hat{\xi}_1) + \tfrac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)\tilde{f}_2(\hat{\xi}_2) + \hat{V}_y(\hat{\xi}, y)\dot{y} + \tfrac{1}{2}(y - \tilde{h}_{21}(\hat{\xi}_1) - \\
&\tilde{h}_{22}(\hat{\xi}_2))^T(W - 4I)(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2)) = 0,\; \hat{V}(0,0) = 0.
\end{aligned} \tag{3.23}$$

But notice from (3.4), with the measurement noise set to zero,

$$\dot{y} = \mathcal{L}_{\tilde{f}_1(\xi_1)+\tilde{g}_{11}(\xi)w}\tilde{h}_{21}(\xi_1) + \mathcal{L}_{\tilde{f}_2(\xi_2)+\tilde{g}_{21}(\xi)w}\tilde{h}_{22}(\xi_2),$$

where $\mathcal{L}$ is the Lie-derivative operator (Vidyasagar, 1993). Then, under certainty-equivalence and with $\mathsf{E}\{w\} = 0$, we have

$$\dot{y} = \mathcal{L}_{\tilde{f}_1(\hat{\xi}_1)}\tilde{h}_{21}(\hat{\xi}_1) + \mathcal{L}_{\tilde{f}_2(\hat{\xi}_2)}\tilde{h}_{22}(\hat{\xi}_2) = \nabla_{\hat{\xi}_1}h_{21}(\hat{\xi}_1)\tilde{f}_1(\hat{\xi}_1) + \nabla_{\hat{\xi}_2}h_{22}(\hat{\xi}_2)\tilde{f}_2(\hat{\xi}_2).$$

Substituting this now in the HJBE (3.23), we have the following HJBE

$$\begin{aligned}\hat{V}_{\hat{\xi}_1}(\hat{\xi},y)\tilde{f}_1(\hat{\xi}_1) + \tfrac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi},y)\tilde{f}_2(\hat{\xi}_2) + \hat{V}_y(\hat{\xi},y)[\nabla_{\hat{\xi}_1}h_{21}(\hat{\xi}_1)\tilde{f}_1(\hat{\xi}_1) + \nabla_{\hat{\xi}_2}h_{22}(\hat{\xi}_2)\tilde{f}_2(\hat{\xi}_2)]+ \\ \tfrac{1}{2}(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))^T(W-4I)(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2)) = 0, \ \hat{V}(0,0) = 0. \quad (3.24)\end{aligned}$$

We then have the following result.

Proposition 3.1.2. *Consider the nonlinear system (3.2) and the $\mathcal{H}_2$ filtering problem for this system. Suppose the plant $\mathbf{P^a_{sp}}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and zero-input observable. Further, suppose there exist a local diffeomorphism φ that transforms the system to the partially decoupled form (3.4), a C^1 positive-semidefinite function $\hat{V} : \hat{N} \times \hat{\Upsilon} \to \Re_+$ locally defined in a neighborhood $\hat{N} \times \hat{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\hat{\xi}, y) = (0,0)$, and matrix functions $\hat{L}_i : \hat{N} \times \hat{\Upsilon} \to \Re^{n_i \times m}$, $i = 1,2$, satisfying the HJBE (3.24) together with the side-conditions (3.20), (3.21). Then the filter $\mathbf{F}^a_{1c}$ solves the $\mathcal{H}_2$ filtering problem for the system locally in $\hat{N}$.*

Proof: The optimality of the filter gains $\hat{L}_1^\star$, $\hat{L}_2^\star$ has already been shown above. It remains to prove asymptotic convergence of the estimation error vector. Accordingly, let $\hat{V} \geq 0$ be a C^1 solution of the HJBE (3.22) or equivalently (3.23). Then, differentiating this solution along a trajectory of (3.16), with $\hat{L}_1 = \hat{L}_1^\star$, $L_2 = \hat{L}_2^\star$, we get

$$\dot{\hat{V}} \;=\; \hat{V}_{\hat{\xi}_1}(\hat{\xi},y)[\tilde{f}_1(\hat{\xi}_1) + \hat{L}_1^\star(\hat{\xi},y)(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))] + \frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi},y)[\tilde{f}_2(\hat{\xi}_2) +$$

$$\begin{aligned}&\hat{L}_2^\star(\hat{\xi},y)(y-\tilde{h}_{21}(\hat{\xi}_1)-\tilde{h}_{22}(\hat{\xi}_2))]+\hat{V}_y(\hat{\xi},y)\dot{y}\\ =\;& -\frac{1}{2}\|z\|_W^2,\end{aligned}$$

where the last equality follows from using the HJBE (3.22). Therefore, the filter dynamics is stable, and $\hat{V}(\hat{\xi},y)$ is non-increasing along a trajectory of (3.16). Further, the condition that $\dot{\hat{V}}(\hat{\xi}(t),y(t))\equiv 0\ \forall t\geq t_s$ implies that $z\equiv 0$, which further implies that $y=\tilde{h}_{21}(\hat{\xi}_1)+\tilde{h}_{22}(\hat{\xi}_2)\ \forall t\geq t_s$. By the zero-input observability of the system, this implies that $\hat{\xi}=\xi$. Finally, since φ is invertible and $\varphi(0)=0$, $\hat{\xi}=\xi$ implies $\hat{x}=\varphi^{-1}(\hat{\xi})=\varphi^{-1}(\xi)=x$. $\square$

Next, we consider the limiting behavior of the filter (3.16) and the corresponding HJBE (3.23). Letting $\varepsilon\downarrow 0$, we obtain from (3.16), and since the vector-field $\tilde{f}_2$ is locally asymptotically stable, we have in the steady-state, the reduced-order filter

$$\bar{\mathbf{F}}^a_{1r}:\left\{\;\dot{\hat{\xi}}_1 \;=\; \tilde{f}_1(\hat{\xi}_1)+\hat{L}_1(\hat{\xi}_1,y)[y-\tilde{h}_{21}(\hat{\xi}_1)]+O(\varepsilon),\right. \tag{3.25}$$

$\hat{\xi}_2\to 0$ and $V_{\hat{\xi}_2}(\hat{\xi},y)\hat{L}_2(\hat{\xi},y)\to 0$. While (3.23) reduces to the Lyapunov-inequality

$$\bar{V}_{\hat{\xi}_2}(\hat{\xi},y)\tilde{f}_2(\hat{\xi},y)\leq 0,\;\;\bar{V}(0,0)=0. \tag{3.26}$$

Note also that, since we are dealing with an infinite-horizon situation, a boundary-layer term does not exist. Moreover, we can then represent the solution of (3.23) locally about $\hat{\xi}=0$ as

$$\hat{V}(\hat{\xi},y)=\bar{V}(\hat{\xi},y)+O(\varepsilon).$$

To relate the above result with the linear theory (Gajic, 1994), (Haddad, 1976), (Haddad, 1977), (Sebald, 1978), we consider the following linear singularly-perturbed system (LSPS):

$$\mathbf{P}^l_{sp}\;:\;\left\{\begin{array}{rcl}\dot{x}_1 &=& A_1x_1+A_{12}x_2+B_{11}w;\quad x_1(t_0)=x_{10}\\ \varepsilon\dot{x}_2 &=& A_{21}x_1+A_2x_2+B_{21}w;\quad x_2(t_0)=x_{20}\\ y &=& C_{21}x_1+C_{22}x_2+w,\end{array}\right. \tag{3.27}$$

where $A_1\in\Re^{n_1\times n_1}$, $A_{12}\in\Re^{n_1\times n_2}$, $A_{21}\in\Re^{n_2\times n_1}$, $A_2\in\Re^{n_2\times n_2}$, $B_{11}\in\Re^{n_1\times m}$, and $B_{21}\in$

$\Re^{n_2 \times m}$, while the other matrices have compatible dimensions. Then, an explicit form of the required transformation φ above is given by the Chang transformation (Chang, 1972):

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} I_{n_1} - \varepsilon \mathsf{HL} & -\varepsilon \mathsf{H} \\ \mathsf{L} & I_{n_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{3.28}$$

where the matrices L and H satisfy the equations

$$\begin{aligned} 0 &= A_2 \mathsf{L} - A_{21} - \varepsilon \mathsf{L}(A_1 - A_{12}\mathsf{L}), \\ &= -\mathsf{H}(A_2 + \varepsilon \mathsf{L} A_{12}) + A_{12} + \varepsilon (A_1 - A_{12}\mathsf{L})\mathsf{H}. \end{aligned}$$

The system is then represented in the new coordinates by

$$\tilde{\mathbf{P}}^l_{sp} : \begin{cases} \dot{\xi}_1 &= \tilde{A}_1 \xi_1 + \tilde{B}_{11} w; \quad \xi_1(t_0) = \xi_{10} \\ \varepsilon \dot{\xi}_2 &= \tilde{A}_2 \xi_2 + \tilde{B}_{21} w; \quad \xi_2(t_0) = \xi_{20} \\ y &= \tilde{C}_{21}\xi_1 + \tilde{C}_{22}\xi_2 + w, \end{cases} \tag{3.29}$$

where

$$\begin{aligned} \tilde{A}_1 &= A_1 - A_{12}\mathsf{L} = A_1 - A_{12}A_2^{-1}A_{21} + O(\varepsilon) \\ \tilde{B}_{11} &= B_{11} - \varepsilon \mathsf{HL} B_{11} - \mathsf{H} B_{21} = B_{11} - A_{12}A_2^{-1}B_{21} + O(\varepsilon) \\ \tilde{A}_2 &= A_2 + \varepsilon \mathsf{L} A_{12} = A_2 + O(\varepsilon) \\ \tilde{B}_{21} &= B_{21} + \varepsilon \mathsf{L} B_{11} = B_{21} + O(\varepsilon) \\ \tilde{C}_{21} &= C_{21} - C_{22}\mathsf{L} = C_{21} - C_{22}A_2^{-1}A_{21} + O(\varepsilon) \\ \tilde{C}_{22} &= C_{22} + \varepsilon (C_{21} - C_{22})\mathsf{H} = C_{22} + O(\varepsilon). \end{aligned}$$

Adapting the filter (3.16) to the system (3.29) yields the following filter

$$\mathbf{F}^l_{1c} : \begin{cases} \dot{\hat{\xi}}_1 &= \tilde{A}_1 \hat{\xi}_1 + \hat{L}_1 (y - \tilde{C}_{21}\hat{\xi}_1 - \tilde{C}_{22}\hat{\xi}_2) \\ \varepsilon \dot{\hat{\xi}}_2 &= \tilde{A}_2 \hat{\xi}_2 + \hat{L}_2 (y - \tilde{C}_{21}\hat{\xi}_1 - \tilde{C}_{22}\hat{\xi}_2). \end{cases} \tag{3.30}$$

Consequently, we have the following Corollary to Proposition 3.1.2. We may assume for simplicity and without loss of generality that the covariance of the noise is $W = I$.

Corollary 3.1.1. *Consider the linear system (3.27) and the $\mathcal{H}_2$ filtering problem for this system. Suppose the plant $\mathbf{P}^l_{sp}$ is asymptotically stable about the equilibrium-point $x = 0$ and observable. Suppose further, it is transformable to the form (3.29), and there exist positive-semidefinite matrices $P_1 \in \Re^{n_1 \times n_1}$, $P_2 \in \Re^{n_2 \times n_2}$, $Q \in \Re^{m \times m}$, and matrices $\hat{L}_1, \hat{L}_2 \in \Re^{n \times m}$, satisfying the linear-matrix-inequalities (LMIs)*

$$\begin{bmatrix} \tilde{A}_1^T \hat{P}_1 + \hat{P}_1 \tilde{A}_1 - 3\tilde{C}_{21}^T \tilde{C}_{21} & -\tilde{C}_{22}^T \tilde{C}_{21} & 3\tilde{C}_{21}^T & 0 \\ -\tilde{C}_{21}^T \tilde{C}_{22} & \frac{1}{\varepsilon}\tilde{A}_2^T \hat{P}_2 + \frac{1}{\varepsilon}\hat{P}_2 \tilde{A}_2 - 3\tilde{C}_{22}^T \tilde{C}_{22} & 3\tilde{C}_{22}^T & 0 \\ 3\tilde{C}_{21} & 3\tilde{C}_{22} & -3I & Q \\ 0 & 0 & Q^T & 0 \end{bmatrix} \le 0 \tag{3.31}$$

$$\begin{bmatrix} 0 & 0 & \frac{1}{2}(\hat{P}_1 L_1 - \tilde{C}_{21}^T) \\ 0 & 0 & -\frac{1}{2}\tilde{C}_{22}^T \\ \frac{1}{2}(\hat{P}_1 L_1 - \tilde{C}_{21}^T)^T & -\frac{1}{2}\tilde{C}_{22}^T & (1-\mu_1)I \end{bmatrix} \le 0 \tag{3.32}$$

$$\begin{bmatrix} 0 & 0 & -\frac{1}{2}\tilde{C}_{21}^T \\ 0 & 0 & \frac{1}{2\varepsilon}(\hat{P}_2 L_2 - \tilde{C}_{22}^T) \\ -\frac{1}{2}\tilde{C}_{21} & \frac{1}{2\varepsilon}(\hat{P}_2 L_2 - \tilde{C}_{22}^T)^T & (1-\mu_2)I \end{bmatrix} \le 0 \tag{3.33}$$

for some numbers $\mu_1, \mu_2 \ge 1$. Then the filter $\mathbf{F}^l_{1c}$ solves the $\mathcal{H}_2$ filtering problem for the system.

Proof: Take

$$\hat{V}(\hat{\xi}, y) = \frac{1}{2}(\hat{\xi}_1^T P_1 \hat{\xi}_1 + \hat{\xi}_2^T P_2 \hat{\xi}_2 + y^T Q y)$$

and apply the result of the Proposition. □

Moreover, the limiting behavior of the filter (3.30) is the reduced-order filter

$$\bar{\mathbf{F}}^l_{1r} : \left\{ \; \dot{\hat{\xi}}_1 \;\; = \;\; \tilde{A}_1 \hat{\xi}_1 + \hat{L}_1 (y - \tilde{C}_{21} \hat{\xi}_1) + O(\varepsilon), \right. \tag{3.34}$$

and the limiting behavior of (3.32) as $\varepsilon \downarrow 0$ is $\bar{P}_2$ satisfies the Lyapunov inequality

$$\tilde{A}_2^T \bar{P}_2 + \bar{P}_2 \tilde{A}_2 \leq 0. \tag{3.35}$$

Proposition 3.1.2 has not yet exploited the benefit of the coordinate transformation in designing the filter (3.16) for the system (3.4). We shall now design separate reduced-order filters for the decomposed subsystems which should be more efficient than the previous one. For this purpose, we let $\varepsilon \downarrow 0$ in (3.4) and obtain the following reduced system model:

$$\widetilde{\mathbf{P}}^{\mathbf{a}}_{\mathbf{r}} : \begin{cases} \dot{\xi}_1 &= \tilde{f}_1(\xi_1) + \tilde{g}_{11}(\xi)w \\ 0 &= \tilde{f}_2(\xi_2) + \tilde{g}_{21}(\xi)w \\ y &= \tilde{h}_{21}(\xi_1) + \tilde{h}_{22}(\xi_2) + \tilde{k}_{21}(\xi)w. \end{cases} \tag{3.36}$$

Then, we assume the following (Khalil, 1985), (Kokotovic, 1986).

Assumption 3.1.1. *The system (3.2), (3.36) is in the "standard form", i.e., the equation*

$$0 = \tilde{f}_2(\xi_2) + \tilde{g}_{21}(\xi)w \tag{3.37}$$

has $l \geq 1$ distinct roots, we can denote any one of these solutions by

$$\bar{\xi}_2 = q(\xi_1, w). \tag{3.38}$$

Under Assumption 3.1.1, we obtain the reduced-order slow subsystem

$$\mathbf{P}^{\mathbf{a}}_{\mathbf{r}} : \begin{cases} \dot{\xi}_1 &= \tilde{f}_1(\xi_1) + \tilde{g}_{11}(\xi_1, q(\xi_1, w))w + O(\varepsilon) \\ y &= \tilde{h}_{21}(\xi_1) + \tilde{h}_{22}(q(\xi_1, w)) + \tilde{k}_{21}(\xi_1, q(\xi_1, w))w + O(\varepsilon) \end{cases} \tag{3.39}$$

and a boundary-layer (or quasi-steady-state) subsystem as

$$\frac{d\bar{\xi}_2}{d\tau} = \tilde{f}_2(\bar{\xi}_2(\tau)) + \tilde{g}_{21}(\xi_1, \bar{\xi}_2(\tau))w, \tag{3.40}$$

where $\tau = t/\varepsilon$ is a stretched-time parameter. It can be shown that there exists $\varepsilon^\star > 0$ such that the subsystem is locally asymptotically stable for all $\varepsilon \geq \varepsilon^\star$ (see Theorem 8.2 in Ref. (Khalil, 1985)) if the original system (3.2) is locally asymptotically stable.

We can therefore proceed to redesign the filter in (3.16) for the composite system (3.39), (3.40) separately as

$$\widetilde{\mathbf{F}}^a_{2c} : \begin{cases} \dot{\breve{\xi}}_1 &= \tilde{f}_1(\breve{\xi}_1) + \breve{L}_1(\breve{\xi}_1, y)[y - \tilde{h}_{21}(\breve{\xi}_1) - h_{22}(q(\breve{\xi}_1, 0))] \\ \varepsilon\dot{\breve{\xi}}_2 &= \tilde{f}_2(\breve{\xi}_2) + \breve{L}_2(\breve{\xi}_2, y)[y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(\breve{\xi}_2)]. \end{cases} \tag{3.41}$$

Notice that ξ_2 cannot be estimated from (3.38) since this is a "quasi-steady-state" approximation.

The following theorem then summaries this design approach.

Theorem 3.1.1. *Consider the nonlinear system (3.2) and the $\mathcal{H}_2$ estimation problem for this system. Suppose the plant $\mathbf{P^a_{sp}}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and zero-input observable. Further, suppose there exist a local diffeomorphism φ that transforms the system to the partially decoupled form (3.4), and Assumption 3.1.1 holds. In addition, suppose there exist C^1 positive-semidefinite functions $\breve{V}_i : \breve{N}_i \times \breve{\Upsilon}_i \to \Re_+,\ i = 1, 2,$ locally defined in neighborhoods $\breve{N}_i \times \breve{\Upsilon}_i \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\breve{\xi}_i, y) = (0, 0),\ i = 1, 2$ respectively, and matrix functions $\breve{L}_i : \breve{N}_i \times \breve{\Upsilon}_i \to \Re^{n_i \times m}$, $i = 1, 2$ satisfying the HJBEs:*

$$\begin{aligned} &\breve{V}_{1\breve{\xi}_1}(\breve{\xi}_1, y)\tilde{f}_1(\breve{\xi}_1) + \breve{V}_{1y}(\breve{\xi}_1, y)[\nabla_{\hat{\xi}_1}\tilde{h}_{21}(\breve{\xi}_1)\tilde{f}_1(\breve{\xi}_1) + \nabla_{\breve{\xi}_1}\tilde{h}_{22}(q(\breve{\xi}_1, 0))\tilde{f}_1(\breve{\xi}_1)] + \tfrac{1}{2}(y - \tilde{h}_{21}(\breve{\xi}_1) - \\ &\qquad \tilde{h}_{22}(q(\breve{\xi}_1, 0))^T(W - 2I)(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(q(\breve{\xi}_1, 0))) = 0,\ \breve{V}_1(0, 0) = 0 \end{aligned} \tag{3.42}$$

$$\begin{aligned} &\tfrac{1}{\varepsilon}\breve{V}_{2\breve{\xi}_2}(\breve{\xi}_2, y)\tilde{f}_2(\breve{\xi}_2) + \breve{V}_{2y}(\breve{\xi}_2, y)[\nabla_{\hat{\xi}_1}\tilde{h}_{21}(\breve{\xi}_1)\tilde{f}_1(\breve{\xi}_1) + \nabla_{\breve{\xi}_2}\tilde{h}_{22}(\breve{\xi}_2)\tilde{f}_2(\breve{\xi}_2)] + \\ &\tfrac{1}{2}(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(\breve{\xi}_2))^T(W - 2I)(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(\breve{\xi}_2)) = 0,\ \breve{V}_2(0, 0) = 0, \end{aligned} \tag{3.43}$$

together with the side-conditions

$$\breve{V}_{1\breve{\xi}_1}(\breve{\xi}_1, y)\breve{L}_1(\breve{\xi}_1, y) = -(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(q(\breve{\xi}_1, 0)))^T, \tag{3.44}$$

$$\breve{V}_{2\breve{\xi}_2}(\breve{\xi}_2, y)\breve{L}_2(\breve{\xi}, y) = -\varepsilon(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(\breve{\xi}_2))^T. \tag{3.45}$$

Then the filter $\widetilde{\mathbf{F}}^{\mathbf{a}}_{\mathbf{2c}}$ *solves the* $\mathcal{H}_2$ *filtering problem for the system locally in* $\cup \breve{N}_i$.

Proof: We define separately two Hamiltonian functions $H_i : T^{\star}\mathcal{X} \times \mathcal{W} \times \Re^{n_i \times m} \to \Re,\ i = 1, 2$ for each of the two separate components of the filter (3.41). Then, the rest of the proof follows along the same lines as Proposition 3.1.2. $\Box$

Again, we have the limiting behavior of filter $\widetilde{\mathbf{F}}^{\mathbf{a}}_{\mathbf{2c}}$ and associated HJBEs (3.42), (3.43) as $\varepsilon \downarrow 0$:

$$\bar{\mathbf{F}}^{a}_{2r} : \begin{cases} \dot{\breve{\xi}}_1 &= \tilde{f}(\breve{\xi}_1) + \breve{L}_1(\breve{\xi}_1, y)[y - \tilde{h}_{21}(\hat{\xi}_1)] \\ \breve{\xi}_2 &\to 0, \end{cases} \tag{3.46}$$

$$\breve{V}_{1\breve{\xi}_1}(\breve{\xi}_1, y)\tilde{f}_1(\breve{\xi}_1) + \breve{V}_{1y}(\breve{\xi}, y)[\nabla_{\breve{\xi}_1}\tilde{h}_{21}(\breve{\xi}_1)\tilde{f}_1(\breve{\xi}_1) + \nabla_{\breve{\xi}_1}\tilde{h}_{22}(q(\breve{\xi}_1, 0))\tilde{f}_1(\breve{\xi}_1)] + \frac{1}{2}(y - \tilde{h}_{21}(\breve{\xi}_1))^T(W - 2I)(y - \tilde{h}_{21}(\breve{\xi}_1)) = 0, \quad \breve{V}_1(0,0) = 0 \tag{3.47}$$

$$\breve{V}_{2\breve{\xi}_2}(\breve{\xi}_2, y)\tilde{f}_2(\breve{\xi}_1) \leq 0, \tag{3.48}$$

together with the side-conditions

$$\breve{V}_{1\breve{\xi}_1}(\breve{\xi}_1, y)\breve{L}_1(\breve{\xi}, y) = -(y - \tilde{h}_{21}(\breve{\xi}_1))^T, \tag{3.49}$$

$$\breve{V}_{2\breve{\xi}_2}(\breve{\xi}, y)\breve{L}_2(\breve{\xi}, y) \to 0. \tag{3.50}$$

Similarly, specializing the result of Theorem 3.1.1 to the linear system (3.27), we obtain the filter

$$\mathbf{F}^{l}_{2c} : \begin{cases} \dot{\breve{\xi}}_1 &= \tilde{A}_1\breve{\xi}_1 + \breve{L}_1(y - \tilde{C}_{21}\breve{\xi}_1) \\ \varepsilon\dot{\breve{\xi}}_2 &= \tilde{A}_2\breve{\xi}_2 + \breve{L}_2(y - \tilde{C}_{21}\breve{\xi}_1 - \tilde{C}_{22}\breve{\xi}_2). \end{cases} \tag{3.51}$$

The following corollary summarizes this development if we assume $W = I$ without loss of generality.

Corollary 3.1.2. *Consider the linear system (3.27) and the* $\mathcal{H}_2$ *filtering problem for this system. Suppose the plant* $\mathbf{P}^{l}_{sp}$ *is asymptotically stable about the equilibrium-point* $x = 0$ *and* *observable. Suppose further, it is transformable to the form (3.29) and A_2 is nonsingular. In addition, suppose there exist positive-semidefinite matrices $\breve{P}_1 \in \Re^{n_1 \times n_1}$, $\breve{P}_2 \in \Re^{n_2 \times n_2}$, $\breve{Q}_1, \breve{Q}_2 \in \Re^{m \times m}$ and matrices $\breve{L}_1 \in \Re^{n_1 \times m}$,$\breve{L}_2 \in \Re^{n_2 \times m}$, satisfying the linear-matrix-inequalities (LMIs)*

$$\begin{bmatrix} \tilde{A}_1^T \breve{P}_1 + \breve{P}_1 \tilde{A}_1 - \tilde{C}_{21}^T \tilde{C}_{21} & \tilde{C}_{21}^T & \breve{P}_1 \\ \tilde{C}_{21}^T & -I & \breve{Q}_1 \\ 0 & \breve{Q}_1 & 0 \end{bmatrix} \leq 0 \tag{3.52}$$

$$\begin{bmatrix} 0 & -\tilde{C}_{21}^T \tilde{C}_{22} & \tilde{C}_{21}^T & 0 \\ -\tilde{C}_{22}^T \tilde{C}_{21} & \frac{1}{\varepsilon}(\tilde{A}_2^T \breve{P}_2 + \breve{P}_2 \tilde{A}_2) - \tilde{C}_{22}^T \tilde{C}_{21} & \tilde{C}_{22}^T & \breve{Q}_2 \\ \tilde{C}_{21} & \tilde{C}_{22} & -I & 0 \\ 0 & \breve{Q}_2 & 0 & 0 \end{bmatrix} \leq 0 \tag{3.53}$$

$$\begin{bmatrix} 0 & 0 & \frac{1}{2}(\breve{P}_1 \breve{L}_1 - \tilde{C}_{21}^T) \\ 0 & 0 & 0 \\ \frac{1}{2}(\breve{P}_1 \breve{L}_1 - \tilde{C}_{21}^T)^T & 0 & (1-\delta_1) I \end{bmatrix} \leq 0 \tag{3.54}$$

$$\begin{bmatrix} 0 & 0 & -\frac{1}{2}\tilde{C}_{21}^T \\ 0 & 0 & \frac{1}{2\varepsilon}(\breve{P}_2 \breve{L}_2 - \tilde{C}_{22}^T) \\ -\frac{1}{2}\tilde{C}_{21} & \frac{1}{2\varepsilon}(\breve{P}_2 \breve{L}_2 - \tilde{C}_{22}^T)^T & (1-\delta_2) I \end{bmatrix} \leq 0 \tag{3.55}$$

for some numbers $\delta_1, \delta_2 \geq 1$. Then the filter $\mathbf{F}_{2c}^l$ solves the $\mathcal{H}_2$ filtering problem for the system.

Proof: Take

$$\begin{aligned} \breve{V}_1(\breve{\xi}_1, y) &= \frac{1}{2}(\breve{\xi}_1^T \breve{P}_1 \breve{\xi}_1 + y^T \breve{Q}_1 y) \\ \breve{V}_2(\breve{\xi}_2, y) &= \frac{1}{2}(\breve{\xi}_2^T \breve{P}_2 \breve{\xi}_2 + y^T \breve{Q}_2 y) \end{aligned}$$

and apply the result of the Theorem. Moreover, the nonsingularity of A_2 guarantees that a reduced-order subsystem exists. $\square$

3.1.3 Aggregate Filters

In the absence of the coordinate transformation, φ, discussed in the previous section, a filter has to be designed to solve the problem for the aggregate system (3.2). We discuss this class of filters in this section. Accordingly, consider the following class of filters:

$$\mathbf{F}^a_{3ag}:\begin{cases}\dot{\grave{x}}_1 &= f_1(\grave{x})+\grave{L}_1(\grave{x},y)[y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2)];\;\;\grave{x}_1(t_0)=\bar{x}_{10}\\ \varepsilon\dot{\grave{x}}_2 &= f_2(\grave{x})+\grave{L}_2(\grave{x},y)(y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2));\;\;\grave{x}_2(t_0)=\bar{x}_{20}\\ \grave{z} &= y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2),\end{cases} \tag{3.56}$$

where $\grave{L}_1 \in \Re^{n_1\times m}$, $\grave{L}_2 \in \Re^{n_2\times m}$ are the filter gains, and $\grave{z}$ is the new penalty variable. We can repeat the same kind of derivation above to arrive at the following result.

Theorem 3.1.2. *Consider the nonlinear system (3.2) and the $\mathcal{H}_2$ estimation problem for this system. Suppose the plant $\mathbf{P^a_{sp}}$ is locally asymptotically stable about the equilibrium-point $x=0$, and zero-input observable. Further, suppose there exists a C^1 positive-semidefinite function $\grave{V}:\grave{N}\times\grave{\Upsilon}\to\Re_+$, locally defined in a neighborhood $\grave{N}\times\grave{\Upsilon}\subset\mathcal{X}\times\mathcal{Y}$ of the origin $(\grave{x}_1,\grave{x}_2,y)=(0,0,0)$, and matrix functions $\grave{L}_i:\grave{N}\times\grave{\Upsilon}\to\Re^{n_i\times m}$, $i=1,2$, satisfying the HJBE:*

$$\begin{aligned}&\grave{V}_{\grave{x}_1}(\grave{x},y)f_1(\grave{x})+\tfrac{1}{\varepsilon}\grave{V}_{\grave{x}_2}(\grave{x},y)f_2(\grave{x})+\grave{V}_y(\grave{x},y)[\nabla_{\grave{x}_1}h_{21}(\grave{x}_1)f_1(\grave{x})+\nabla_{\grave{x}_2}h_{22}(\grave{x}_2)f_2(\grave{x})]+\\ &\tfrac{1}{2}(y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2))^T(W-2I)(y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2))=0,\;\grave{V}(0,0)=0\end{aligned} \tag{3.57}$$

together with the side-conditions

$$\grave{V}_{\grave{x}_1}(\grave{x},y)\grave{L}_1(\grave{x},y) = -(y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2))^T \tag{3.58}$$

$$\grave{V}_{\grave{x}_2}(\grave{x},y)\grave{L}_2(\grave{x},y) = -\varepsilon(y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2))^T. \tag{3.59}$$

Then, the filter $\mathbf{F}^a_{3ag}$ solves the $\mathcal{H}_2$-filtering problem for the system locally in $\grave{N}$.

Proof: Proof follows along the same lines as Proposition 3.1.2.

The result of Theorem 3.1.2 can similarly be specialized to the linear system $\mathbf{P}^l_{sp}$. Moreover, we also have the limiting behavior of the filter $\mathbf{F}^a_{3ag}$ as $\varepsilon \downarrow 0$

$$\bar{\mathbf{F}}^a_{3ag} : \begin{cases} \dot{\grave{x}}_1 &= f_1(\grave{x}) + \grave{L}_1(\grave{x}, y)[y - h_{21}(\grave{x}_1)]; \;\; \grave{x}_1(t_0) = \bar{x}_{10} \\ \grave{x}_2 &\to 0 \end{cases} \tag{3.60}$$

$$\grave{V}_{\grave{x}_2}(\grave{x}, y) f_2(\grave{x}) \leq 0, \tag{3.61}$$

together with the side-conditions

$$\grave{V}_{\grave{x}_1}(\grave{x}, y)\grave{L}_1(\grave{x}, y) = -(y - h_{21}(\grave{x}_1))^T \tag{3.62}$$

$$\grave{V}_{\grave{x}_2}(\grave{x}, y)\grave{L}_2(\grave{x}, y) \to 0. \tag{3.63}$$

Remark 3.1.2. *Also, comparing the accuracy of the filters* $\mathbf{F}^a_{1c}$*,* $\mathbf{F}^a_{2c}$*,* $\mathbf{F}^a_{3ga}$*, we see that the order of the accuracy is* $\mathbf{F}^a_{2c} \succeq \mathbf{F}^a_{1c} \succeq \mathbf{F}^a_{3ag}$ *by virtue of the decomposition, where the relational operator "*$\succeq$*" implies better.*

3.1.4 Push-Pull Configuration

Finally, in this subsection, we present a *"push-pull"* configuration for the aggregate filter presented in the above section. Since the dynamics of the second subsystem is fast, we can afford to reduce the gain of the filter for this subsystem to avoid instability, while for the slow subsystem, we can afford to increase the gain. Therefore, we consider the following filter configuration

$$\mathbf{F}^a_{4ag} : \begin{cases} \dot{x}^\flat_1 &= f_1(x^\flat) + (L^\flat_1 + L^\flat_2)(x^\flat, y)[y - h_{21}(x^\flat_1) - h_{22}(x^\flat_2)]; \; x^\flat_1(t_0) = \bar{x}_{10} \\ \varepsilon\dot{x}^\flat_2 &= f_2(x^\flat) + (L^\flat_1 - L^\flat_2)(x^\flat, y)[y - h_{21}(x^\flat_1) - h_{22}(x^\flat_2)]; \; x^\flat_2(t_0) = \bar{x}_{20} \\ z^\flat &= y - h_{21}(x^\flat_1) - h_{22}(x^\flat_2), \end{cases} \tag{3.64}$$

where $x^\flat \in \mathcal{X}$ is the filter state, $L^\flat_1 \in \Re^{n_1 \times m}$, $L^\flat_2 \in \Re^{n_2 \times m}$ are the filter gains, while all the other variables have their corresponding previous meanings and dimensions.

Consequently, going through similar manipulations as in Proposition 3.1.2 we can arrive at the following result.

Proposition 3.1.3. *Consider the nonlinear system (3.2) and the $\mathcal{H}_2$ estimation problem for this system. Suppose the plant $\mathbf{P^a_{sp}}$ is locally asymptotically stable about the equilibrium-point $x = 0$, and zero-input observable. Further, suppose there exist a C^1 positive-semidefinite function $V^\flat : N^\flat \times \Upsilon^\flat \to \Re_+$, locally defined in a neighborhood $N^\flat \times \Upsilon^\flat \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(x_1^\flat, x_2^\flat, y) = (0, 0, 0)$, and matrix functions $L_1^\flat \in \Re^{n_1 \times m}$, $L_2^\flat :\in \Re^{n_2 \times m}$, satisfying the HJBE:*

$$
\begin{aligned}
V^\flat_{x_1^\flat}(x^\flat, y) f_1(x^\flat) + \tfrac{1}{\varepsilon} V^\flat_{x_2^\flat}(x^\flat, y) f_2(x_1^\flat) + V^\flat_y(x^\flat, y)[\nabla_{x_1^\flat} h_{21}(x_1^\flat) f_1(x^\flat) + \nabla_{x_2^\flat} h_{22}(x_2^\flat) f_2(x^\flat)] + \\
\tfrac{1}{2}(y - h_{21}(x_1^\flat) - h_{22}(x_2^\flat))^T (W - 2I)(y - h_{21}(x_1^\flat) - h_{22}(x_2^\flat)) = 0, \quad V^\flat(0, 0) = 0
\end{aligned} \tag{3.65}
$$

together with the side-conditions

$$
(V^\flat_{x_1^\flat} + V^\flat_{x_2^\flat})(x^\flat, y) L_1^\flat(x^\flat, y) \;=\; -(y - h_{21}(x_1^\flat) - h_{22}(x_2^\flat))^T \tag{3.66}
$$

$$
(V^\flat_{x_1^\flat} - V^\flat_{x_2^\flat})(x^\flat, y) L_2^\flat(x^\flat, y) \;=\; -\varepsilon(y - h_{21}(x_1^\flat) - h_{22}(x_2^\flat))^T. \tag{3.67}
$$

Then, the filter $\mathbf{F}^a_{4ag}$ solves the $\mathcal{H}_2$ filtering problem for the system localy in $N^\flat$.

Remark 3.1.3. *If the nonlinear system (3.2) is in the standard form, i.e., the equivalent of Assumption 3.1.1 is satisfied, and there exists at least one root $\bar{x}_2 = \sigma(x_1, w)$ to the equation*

$$
0 = f_2(x_1, x_2) + g_{21}(x_1, x_2) w,
$$

then reduced-order filters can also be constructed for the system similar to the result of Section 3.2.3 and Theorem 3.1.1. Such filters would take the following form

$$
\mathbf{F}^a_{5agr} : \begin{cases}
\dot{\check{x}}_1 = f_1(\check{x}_1, \sigma(\check{x}_1, 0)) + \check{L}_1(\check{x}_1, y)(y - h_{21}(\check{x}_1) - h_{22}(\sigma(\check{x}_1, 0))); \quad \check{x}_1(t_0) = \bar{x}_{10} \\
\varepsilon \dot{\check{x}}_2 = f_2(\check{x}) + \check{L}_2(\check{x}, y)(y - h_{21}(\check{x}_1) - h_{22}(\check{x}_2)); \quad \check{x}_2(t_0) = \bar{x}_{20} \\
\check{z} = y - h_{21}(\check{x}_1) - h_{22}(\check{x}_2).
\end{cases} \tag{3.68}
$$

However, these filters would fall into the class of decomposition filters, rather than aggregate, and because of this, we shall not discuss them further in this section.

In the next section, we consider some examples.

3.1.5 Examples

We consider a few simple examples in this subsection. The first example is designed to illustrate how the use of a transformation and a decomposition of the system can simplify the filter design.

Example 3.1.2. *We reconsider Example 3.1.1 to design a decomposed filter for the system. The system is transformed to the following two systems locally over U_1 and U_2 respectively:*

$$\begin{aligned}\dot{\xi}_1 &= -\xi_1^{-1} + (\frac{1}{\xi_2} - \xi_1)w_0 \\ \varepsilon\dot{\xi}_2 &= -\xi_2 - \xi_2 w_0 \\ y &= -\xi_1\xi_2 - \xi_2 + w_0.\end{aligned}$$

defined on $\widetilde{U}_1 = \Re^2 \setminus \{\xi_2 = 0\}$, and

$$\begin{aligned}\dot{\xi}_1 &= -\xi_1 + w_0 \\ \varepsilon\dot{\xi}_2 &= 0 \\ y &= \xi_1 + w_0.\end{aligned}$$

which is defined for $\widetilde{U}_2 = \{\xi_2 = 0\}$. We design the filter (3.16) for each of the subsystems. Accordingly, it can be checked that, the functions $\hat{V}_1(\hat{\xi}) = \frac{1}{2}(\hat{\xi}_1^{-2} + \varepsilon\hat{\xi}_2^2)$, $\hat{V}_2(\hat{\xi}) = \frac{1}{2}(\hat{\xi}_1^2 + \varepsilon\hat{\xi}_2^2)$ solve the HJBE (3.23) for the filter and for the two subsystems respectively. Therefore, the filter gains can be calculated from (3.20), (3.21) for the two subsystems respectively as

$$\hat{L}_1(\hat{\xi}, y) = \hat{\xi}_1^3(y + \hat{\xi}_1\hat{\xi}_2 + \hat{\xi}_2), \quad \hat{L}_2(\hat{\xi}, y) = -\frac{(y + \hat{\xi}_1\hat{\xi}_2 + \hat{\xi}_2)}{\hat{\xi}_2},$$

and

$$\hat{L}_1(\hat{\xi}, y) = -\frac{(y - \hat{\xi}_1)}{\hat{\xi}_1}, \quad \hat{L}_2(\hat{\xi}, y) = 0.$$

Moreover, the gains $\hat{L}_1$, $\hat{L}_2$ are set to zero if $|\hat{\xi}_1| < \epsilon$, $|\hat{\xi}_2| < \epsilon$ (small) respectively to avoid the singularity at the origin $\hat{\xi}_1 = \hat{\xi}_2 = 0$.

Example 3.1.3. *Consider now the following singularly perturbed nonlinear system*

$$\begin{aligned} \dot{x}_1 &= -x_1^3 + x_2 + x_1^2 w_0 \\ \varepsilon\dot{x}_2 &= -x_1 - x_2 + sin(x_2) w_0 \\ y &= x_1 + x_2 + w_0. \end{aligned}$$

where w_0 is a zero-mean Gaussian white noise process with $W = I$, $\varepsilon \geq 0$. We construct the aggregate filter $\mathbf{F}^a_{3ag}$ presented in the previous section for the above system. It can be checked that the system is locally observable, and the function $\grave{V}(\grave{x}) = \frac{1}{2}(\grave{x}_1^2 + \varepsilon\grave{x}_2^2)$, solves the inequality form of the HJBE (3.57) corresponding to the system. Subsequently, we calculate the gains of the filter as

$$\grave{L}_1(\grave{x}, y) = -\frac{(y - \grave{x}_1 - \grave{x}_2)}{\grave{x}_1}, \quad \grave{L}_2(\grave{x}, y) = -\frac{\varepsilon(y - \grave{x}_1 - \grave{x}_2)}{\grave{x}_2}, \tag{3.69}$$

where again the gains $\grave{L}_1$, $\grave{L}_2$ are set equal to zero if $|\grave{x}_1| < \epsilon$ (small), $|\grave{x}_2| < \epsilon$ (small) to avoid the singularity at $\grave{x}_1 = \grave{x}_2 = 0$.

Similarly, we can construct the push-pull filter gains for the above system since the HJBEs (3.57) and (3.65) are identical as

$$L_1^\flat(x^\flat, y) = -\frac{(y - x_1^\flat - x_2^\flat)}{(x_1^\flat + x_2^\flat)}, \quad L_2^\flat(x^\flat, y) = -\frac{\varepsilon(y - x_1^\flat - x_2^\flat)}{(x_1^\flat + x_2^\flat)}, \tag{3.70}$$

and again, we set the gains $L_1^\flat$, $L_2^\flat$ to zero if $|x_1^\flat + x_2^\flat| < \epsilon$ (small) to avoid the singularity at $x_1^\flat + x_2^\flat = 0$.

3.2 $\mathcal{H}_2$ Filtering for Discrete-time Systems

In this section, we discuss the counterpart discrete-time results for the $\mathcal{H}_2$ local filtering problem for affine nonlinear singularly-perturbed systems. Two types of filters, namely, (i) decomposition, and (ii) aggregate filters will similarly be considered, and sufficient conditions for the solvability of the problem in terms of discrete-time Hamilton-Jacobi-Isaacs equations (DHJIEs) will be presented. We begin with the problem definition and other preliminaries.

3.2.1 Problem Definition and Preliminaries

Figure 3.2 shows the equivalent set-up for the disrete-time problem, where $\mathbf{P}_k$ is the plant, while $\mathbf{F}_k$ is the filter. The noise signal $w_0 \in \mathcal{S}'$ is similarly a bounded spectral signal (e.g. a Gaussian white-noise signal) which belongs to the set $\mathcal{S}'$ of bounded spectral signals, while $\tilde{z} \in \mathcal{P}'$, is a bounded power signal or ℓ_2 signal, which belongs to the space of bounded power signals. Thus, the induced norm from w_0 to $\tilde{z}$ (the penalty variable to be defined later) is the ℓ_2-norm of the interconnected system $\mathbf{F}_k \circ \mathbf{P}_k$, i.e.,

$$\|\mathbf{F}_k \circ \mathbf{P}_k\|_{\ell_2} \stackrel{\Delta}{=} \sup_{0\neq w_0 \in \mathcal{S}'} \frac{\|\tilde{z}\|_{\mathcal{P}'}}{\|w_0\|_{\mathcal{S}'}}, \tag{3.71}$$

where "$\circ$" denote operator composition,

$$\begin{aligned}
\mathcal{P}' &\stackrel{\Delta}{=} \{w:\ w \in \ell_\infty,\ R_{ww}(k),\ S_{ww}(j\omega) \text{ exist for all } k \text{ and all } \omega \text{ resp., } \|w\|_{\mathcal{P}'} < \infty\} \\
\mathcal{S}' &\stackrel{\Delta}{=} \{w:\ w \in \ell_\infty,\ R_{ww}(k),\ S_{ww}(j\omega) \text{ exist for all } k \text{ and all } \omega \text{ resp., } \|S_{ww}(j\omega)\|_\infty < \infty\}
\end{aligned}$$

$$\|z\|^2_{\mathcal{P}'} \stackrel{\Delta}{=} \lim_{K\to\infty} \frac{1}{2K} \sum_{k=-K}^{K} \|z_k\|^2$$

$$\|w_0\|_{\mathcal{S}'} = \sqrt{\|S_{w_0 w_0}(j\omega)\|_\infty} = \sqrt{\sup_w \|S_{w_0 w_0}(j\omega)\|},$$

and R_{ww}, $S_{ww}(j\omega))$ are the autocorrelation and power spectral-density matrices of w. Notice also that, $\|(.)\|_{\mathcal{P}'}$ and $\|(.)\|_{\mathcal{S}'}$ are seminorms. In addition, if the plant is stable, we replace the induced ℓ_2-norm above by the equivalent $\mathcal{H}_2$-subspace norms.

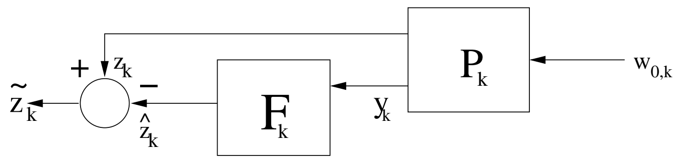

Figure 3.2 Set-up for discrete-time $\mathcal{H}_2$ filtering

At the outset, we consider the following singularly-perturbed affine nonlinear causal discrete-time state-space model of the plant which is defined on $\mathcal{X} \subseteq \Re^{n_1+n_2}$ with zero control input:

$$\mathbf{P}^{da}_{sp} : \begin{cases} x_{1,k+1} &= f_1(x_{1,k}, x_{2,k}) + g_{11}(x_{1,k}, x_{2,k})w_k; \quad x_1(k_0, \varepsilon) = x^{10} \\ \varepsilon x_{2,k+1} &= f_2(x_{1,k}, x_{2,k}, \varepsilon) + g_{21}(x_{1,k}, x_{2,k})w_k; \quad x_2(k_0, \varepsilon) = x^{20} \\ y_k &= h_{21}(x_{1,k}) + h_{22}(x_{2,k}) + k_{21}(x_{1,k}, x_{2,k})w_k, \end{cases} \tag{3.72}$$

where $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathcal{X}$ is the state vector with x_1 the slow state which is n_1-dimensional and x_2 the fast, which is n_2-dimensional; $w \in \mathcal{W} \subseteq \Re^r$ is an unknown disturbance (or noise) signal, which belongs to the set $\mathcal{W}$ of admissible exogenous inputs; $y \in \mathcal{Y} \subset \Re^m$ is the measured output (or observation) of the system, and belongs to $\mathcal{Y}$, the set of admissible measured-outputs; while ε is a small perturbation parameter.

The functions $f_1 : \mathcal{X} \to \mathcal{X} \subseteq \Re^{n_1}$, $f_2 : \mathcal{X} \times \Re \to \mathcal{X}$, $g_{11} : \mathcal{X} \to \mathcal{M}^{n_1 \times m}(\mathcal{X})$, $g_{21} : \mathcal{X} \to \mathcal{M}^{n_2 \times m}(\mathcal{X})$, where $\mathcal{M}^{i \times j}$ is the ring of $i \times j$ smooth matrices over $\mathcal{X}$, $h_{21}, h_{22} : \mathcal{X} \to \Re^m$, and $k_{21} : \mathcal{X} \to \mathcal{M}^{m \times m}(\mathcal{X})$ are real C^∞ functions of x. More explicitly, f_2 is of the form $f_2(x_{1,k}, x_{2,k}, \varepsilon) = (\varepsilon x_{2,k} + \bar{f}_2(x_{1,k}, x_{2,k}))$ for some function $\bar{f} : \mathcal{X} \to \Re^{n_2}$. Furthermore, we assume without any loss of generality that the system (3.72) has an isolated equilibrium-point at $(x_1^T, x_2^T) = (0, 0)$ such that $f_1(0, 0) = 0$, $f_2(0, 0) = 0$, $h_{21}(0, 0) = h_{22}(0, 0) = 0$. We also assume that there exists a unique solution $x(k, k_0, x_0, w, \varepsilon)$ $\forall k \in \mathbf{Z}$ for the system, for all initial conditions $x(k_0) \stackrel{\Delta}{=} x_0 = (x^{10^T},\ x^{20^T})^T$, for all $w \in \mathcal{W}$, and all $\varepsilon \in \Re$.

The standard discrete-time $\mathcal{H}_2$ local filtering/state estimation problem is defined as follows.

Definition 3.2.1. *(Standard $\mathcal{H}_2$ Local State Estimation (Filtering) Problem). Find a filter,* $\mathbf{F}_k$*, for estimating the state x_k or a function of it, $z_k = h_1(x_k)$, from observations $\mathsf{Y}_k \stackrel{\Delta}{=} \{y_i :$*

$i \leq k\}$ of y_i up to time k, to obtain the estimate

$$\hat{x}_k = \mathbf{F}_k(\mathsf{Y}_k),$$

such that, the $\mathcal{H}_2$-norm from the input w to some suitable penalty function $\tilde{z}$ is locally minimized for all initial conditions $x^0 \in \mathcal{O} \subset \mathcal{X}$, for all $w \in \mathcal{W} \subset \mathcal{S}'$. Moreover, if the filter solves the problem for all $x^0 \in \mathcal{X}$, we say the problem is solved globally.

We shall adopt the following definition of local observability.

Definition 3.2.2. *For the nonlinear system $\mathbf{P}_{sp}^{da}$, we say that it is locally zero-input observable, if for all states $x_1,\ x_2 \in U \subset \mathcal{X}$ and input $w(.) = 0$,*

$$y(k; x_1, w) \equiv y(k; x_2, w) \Longrightarrow x_1 = x_2,$$

where $y(., x_i, w), i = 1, 2$ is the output of the system with the initial condition $x_{k_0} = x_i$. Moreover, the system is said to be zero-input observable if it is locally observable at each $x^0 \in \mathcal{X}$ or $U = \mathcal{X}$.

3.2.2 Discrete Decomposition Filters

In this section, we present a decomposition approach to the $\mathcal{H}_2$ estimation problem defined in the previous section, while in the next section, we present an aggregate approach. For this purpose, we assume that the noise signal $w \in \mathcal{W} \subset \mathcal{S}'$ is a zero-mean Gaussian white-noise process, i.e.,

$$\mathsf{E}\{w_k\} = 0, \quad \mathsf{E}\{w_k w_j^T\} = W\delta_{kj}.$$

Also, the initial conditions $x_1(k_0, \varepsilon) = x^{10}$, $x_2(k_0, \varepsilon) = x^{20}$ are assumed to be Gaussian distributed random variables with means given by

$$\mathsf{E}\{x^{10}\} = \bar{x}^{10}, \quad \mathsf{E}\{x^{20}\} = \bar{x}^{20}.$$

We construct two-time scale filters corresponding to the decomposition of the system into a *"fast"* and a *"slow"* subsystems. As in the linear case (Chang, 1972), (Aganovic, 1996), (Kim 2002), (Lim, 1996), (Sadjadi, 1990), we first similarly assume that there exists locally a smooth invertible coordinate transformation (a diffeomorphism) $\varphi : x \mapsto \xi$, i.e.,

$$\xi_1 = \varphi_1(x, \varepsilon), \quad \varphi_1(0, \varepsilon) = 0, \quad \xi_2 = \varphi_2(x, \varepsilon), \quad \varphi_2(0, \varepsilon) = 0, \;\; \xi_1 \in \Re^{n_1}, \xi_2 \in \Re^{n_2}, \tag{3.73}$$

such that the system (3.72) is locally decomposed into the form

$$\tilde{\mathbf{P}}^{\mathbf{da}}_{sp} : \begin{cases} \xi_{1,k+1} &= \tilde{f}_1(\xi_{1,k}, \varepsilon) + \tilde{g}_{11}(\xi_k, \varepsilon) w_k, \quad \xi_1(k_0) = \varphi_1(x^0, \varepsilon) \\ \varepsilon \xi_{2,k+1} &= \tilde{f}_2(\xi_{2,k}, \varepsilon) + \tilde{g}_{21}(\xi_k, \varepsilon) w_k; \quad \xi_2(k_0) = \varphi_2(x^0, \varepsilon) \\ y_k &= \tilde{h}_{21}(\xi_{1,k}, \xi_{2,k}, \varepsilon) + \tilde{h}_{22}(\xi_{1,k}, \xi_{2,k}, \varepsilon) + \tilde{k}_{21}(\xi_k, \varepsilon) w. \end{cases} \tag{3.74}$$

Necessary conditions that such a transformation must satisfy are given in Proposition 3.1.1. A local version of that result can be derived for the discrete-time case. In most cases, this local version would also be sufficient. Moreover, if such a coordinate transformation exists, the functions $\tilde{f}_1$, $\tilde{f}_2$, $\tilde{g}_{11}$, $\tilde{g}_{21}$ will be some nonlinear functions of f_1,f_2, g_{11}, g_{21} since

$$\begin{aligned} \xi_{1,k+1} &= \varphi_1 \begin{pmatrix} f_1(x_k) + g_{11}(x_k) w_k \\ f_2(x_k, \varepsilon) + \tilde{g}_{21}(x_k) w_k, \end{pmatrix} \circ \varphi^{-1}(\xi_k, \varepsilon), \quad \xi_1(k_0) = \varphi_1(x^0, \varepsilon) \\ \varepsilon \xi_{2,k+1} &= \varphi_2 \begin{pmatrix} f_1(x_k) + g_{11}(x_k) w_k \\ f_2(x_k, \varepsilon) + \tilde{g}_{21}(x_k) w_k, \end{pmatrix} \circ \varphi^{-1}(\xi_k, \varepsilon), \quad \xi_2(k_0) = \varphi_2(x^0, \varepsilon) \\ y_k &= h_{21}(\tilde{\varphi}_1(\xi_k, \varepsilon)) + h_{22}(\tilde{\varphi}_2(\xi_k, \varepsilon)) + k_{21}(\varphi^{-1}(\xi_k, \varepsilon)) w_k. \end{aligned}$$

where $\tilde{\varphi}_i = \Pi_i \circ \varphi^{-1}, i =, 1, 2$, $\Pi_i : x \mapsto x_i$ is the natural projection onto the i-th coordinate, and $x^0 = (x^{10T}, x^{20T})^T$.

The following result gives local conditions that φ must satisfy.

Proposition 3.2.1. *Consider the nonlinear system (3.72) defined on $\mathcal{X}$. Let (U_1, x), (U_2, ξ), $U_1, U_2 \subset \mathcal{X}$ containing the origin, be two coordinate neighborhoods on $\mathcal{X}$, and consider the* *problem of finding a local diffeomorphism*[3] $\varphi : U_1 \to U_2$, $\xi = \varphi(x, \varepsilon)$ *so that the system is transformed into the partially decoupled form (3.74) by this coordinate change. Then, the necessary conditions that such a transformation must satisfy are given by the following:*

(i) φ_* *is locally an isomorphism;*

(ii)

$$\left\langle \frac{\partial}{\partial \xi_j}, d\left(\varphi^{-1*}\begin{pmatrix} f_1 \\ f_2 \end{pmatrix}^* \varphi_i\right)\right\rangle = \left\langle \frac{\partial}{\partial \xi_j}, d\left(\left(\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \circ \varphi^{-1}\right)^* \varphi_i\right)\right\rangle = 0, \quad (3.75)$$

$i, j = 1, 2$, $i \neq j$, *and where "$(_*)$", "$(^*)$" are the push-forward and pull-back operators (Boothby, 1975) respectively.*

Proof: Conditions (i), (ii) can be rewritten respectively as

$$det \begin{bmatrix} \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial x_2} \\ \frac{\partial \varphi_2}{\partial x_1} & \frac{\partial \varphi_2}{\partial x_2} \end{bmatrix}(0) \neq 0 \quad (3.76)$$

$$\frac{\partial}{\partial \xi_2}\left(\varphi_1\begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} \circ \varphi^{-1}(\xi, \varepsilon)\right) = 0 \quad (3.77)$$

$$\frac{\partial}{\partial \xi_1}\left(\varphi_2\begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} \circ \varphi^{-1}(\xi, \varepsilon)\right) = 0 \quad (3.78)$$

Then, equation (3.76) which corresponds to (i), guarantees that the transformation φ is locally invertible and satisfies the conditions of the inverse-function Theorem (Sastry, 1999). While equations (3.77) and (3.9) which correspond to (ii) follow from (3.75), (3.75) respectively by setting $w = 0$, and guarantee that $\tilde{f}_1(\xi_1)$, $\tilde{f}_2(\xi_2)$ are independent of ξ_2, ξ_1 respectively. □

Remark 3.2.1. *It is virtually impossible to find a coordinate transformation such that* $\tilde{h}_{2j} = \tilde{h}_{2j}(\xi_j), j = 1, 2$. *Thus, we have made the more practical assumption that* $\tilde{h}_{2j} = \tilde{h}_{2j}(\xi_1, \xi_2), j = 1, 2$.

[3] *see Ref. (Boothby, 1975) for most of the terminology here.*

The filter is then designed based on this transformed model with the optimal noise level set as $w^\star = \mathsf{E}\{w\} = 0$, and accordingly, we propose the following composite filter

$$\mathbf{F}^{da}_{1c} : \begin{cases} \hat{\xi}_{1,k+1} &= \tilde{f}_1(\hat{\xi}_{1,k}, \varepsilon) + L_1(\hat{\xi}_k, y_k, \varepsilon)[y_k - \tilde{h}_{21}(\hat{\xi}_k, \varepsilon) - \tilde{h}_{22}(\hat{\xi}_k, \varepsilon)]; \\ & \hat{\xi}_1(k_0) = \mathsf{E}\{\varphi_1(x^0, \varepsilon)\} \\ \varepsilon\hat{\xi}_{2,k+1} &= \tilde{f}_2(\hat{\xi}_{2,k}, \varepsilon) + L_2(\hat{\xi}_k, y_k, \varepsilon)[y_k - \tilde{h}_{21}(\hat{\xi}_k, \varepsilon) - \tilde{h}_{22}(\hat{\xi}_k, \varepsilon)]; \\ & \hat{\xi}_2(k_0) = \mathsf{E}\{\varphi_2(x^0, \varepsilon)\}. \end{cases} \tag{3.79}$$

where $\hat{\xi} \in \mathcal{X}$ is the filter state, $L_1 \in \Re^{n_1 \times m}$, $L_2 \in \Re^{n_2 \times m}$ are the filter gains, while all the other variables have their corresponding previous meanings and dimensions. We can then define the penalty variable or estimation error at each instant k as

$$\tilde{z}_k = y_k - \tilde{h}_{21}(\hat{\xi}_k) - \tilde{h}_{22}(\hat{\xi}_k). \tag{3.80}$$

The problem can then be similarly formulated as a dynamic optimization problem with the following cost functional

$$\begin{aligned} \min_{\substack{L_1 \in \Re^{n \times m}, L_2 \in \Re^{n_2 \times m}, \\ w \in \mathcal{S}', \hat{\xi}(k_0) = 0}} J_1(L_1, L_2, w) &= \mathsf{E}\left\{\frac{1}{2}\sum_{k_0}^{\infty} \|\tilde{z}_k\|^2_W\right\} = \frac{1}{2}\left\{\|\mathbf{F}^{da}_{1c} \circ \tilde{\mathbf{P}}^{\mathbf{da}}_{sp}\|^2_{\mathcal{H}_2}\right\}_W \\ & s.t. \;\; (3.79) \;\text{ and with } w = 0, \;\; \lim_{k \to \infty}\{\hat{\xi}_k - \xi_k\} = 0; \end{aligned} \tag{3.81}$$

Therefore, to solve it, we consider the expected Hamiltonian function defined as $H : \mathcal{X} \times \mathcal{Y} \times \Re^{n_1 \times m} \times \Re^{n_2 \times m} \times \Re \to \Re$:

$$\begin{aligned} H(\hat{\xi}, y, L_1, L_2, V, \varepsilon) &= V\Big(\tilde{f}_1(\hat{\xi}_1, \varepsilon) + L_1(\hat{\xi}, y, \varepsilon)(y - \tilde{h}_{21}(\hat{\xi}_1, \varepsilon) - h_{22}(\hat{\xi}_2, \varepsilon)), \\ & \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon) + \frac{1}{\varepsilon}L_2(\hat{\xi}, y, \varepsilon)(y - \tilde{h}_{21}(\hat{\xi}, \varepsilon) - \tilde{h}_{22}(\hat{\xi}, \varepsilon)), y\Big) - \\ & V(\hat{\xi}, y_{k-1}) + \frac{1}{2}\|\tilde{z}\|^2_W \end{aligned} \tag{3.82}$$

for some C^1 positive-semidefinite function $V : \mathcal{X} \times \mathcal{Y} \to \Re_+$ and where $\hat{\xi}_1 = \hat{\xi}_{1,k}$, $\hat{\xi}_2 = \hat{\xi}_{2,k}$ $y = y_k$, $z = z_k$. Then the optimal gains $L_1^\star$ and $L_2^\star$ can be obtained by minimizing $H(.,.,L_1,L_2,.,.)$ with respect to L_1 and L_2 in the above expression (3.82), as

$$[L_1^\star, \quad L_2^\star] = \arg\min_{L_1,L_2} H(\hat{\xi}, y, L_1, L_2, V, \varepsilon). \tag{3.83}$$

Because the Hamiltonian function (3.82) is not a linear or quadratic function of the gains L_1, L_2, only implicit solutions can be obtained by solving the equations

$$\begin{aligned} \left.\frac{\partial V(\lambda,\mu,y)}{\partial \lambda}\right|_{\lambda=\lambda^\star,\mu=\mu^\star} &= 0 \\ \left.\frac{\partial V(\lambda,\mu,y)}{\partial \mu}\right|_{\lambda=\lambda^\star,\mu=\mu^\star} &= 0 \end{aligned}$$

for $L_1^\star(\hat{\xi}, y)$, $L_2^\star(\hat{\xi}, y)$ simultaneously, where

$$\lambda = \tilde{f}_1(\hat{\xi}_1) + L_1(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}_1) - h_{22}(\hat{\xi}_2)),$$

$$\mu = \frac{1}{\varepsilon}(\tilde{f}_1(\hat{\xi}_1) + L_2(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}_1) - h_{22}(\hat{\xi}_2))),$$

$\partial V/\partial \lambda$, $\partial V/\partial \mu$ are the row vectors of first-order partial derivatives of V with respect to λ and μ respectively, and V solves the discrete-time Hamilton-Jacobi-Bellman equation (DHJBE)

$$\hat{H}(\hat{\xi}, y, L_1^\star, L_2^\star, V, \varepsilon) = 0, \quad V(0,0,0) = 0, \tag{3.84}$$

with

$$\left.\frac{\partial^2 V}{\partial \lambda^2}\right|_{\lambda=\lambda^\star,\mu=\mu^\star} \geq 0, \quad \left.\frac{\partial^2 V}{\partial \mu^2}\right|_{\lambda=\lambda^\star,\mu=\mu^\star} \geq 0.$$

Thus, the only way to obtain an explicit solution is to use an approximate scheme. Intuitively, a first-order Taylor series expansion of the Hamiltonian about $(\tilde{f}_1(\hat{\xi}_1),$ $\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2), y)$ in the direction of the state vectors $(\hat{\xi}_1, \hat{\xi}_2)$, would capture most if not all of the system dynamics. This approximate Hamiltonian is then given by

$$\begin{aligned}\widehat{H}(\hat{\xi}, y, \hat{L}_1, \hat{L}_2, \hat{V}, \varepsilon) &= \hat{V}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y) - \hat{V}(\hat{\xi}, y_{k-1}) + \\ &\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y)\hat{L}_1(\hat{\xi}, y, \varepsilon)[y - \tilde{h}_{21}(\hat{\xi}, \varepsilon) - h_{22}(\hat{\xi}, \varepsilon)] + \\ &\frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y)\hat{L}_2(\hat{\xi}, y, \varepsilon)[y - \tilde{h}_{21}(\hat{\xi}, \varepsilon) - h_{22}(\hat{\xi}, \varepsilon)] + \\ &\frac{1}{2}\|\tilde{z}\|^2_W + O(\|\hat{\xi}\|^2), \qquad (3.85)\end{aligned}$$

where $\hat{V}$, $\hat{L}_1$, $\hat{L}_2$ are the corresponding approximate functions, and $\hat{V}_{\hat{\xi}_1}$, $\hat{V}_{\hat{\xi}_2}$ are the row vectors of first-partial derivatives of $\hat{V}$ with respect to $\hat{\xi}_1$, $\hat{\xi}_2$ respectively. Completing the squares now for $\hat{L}_1(\hat{\xi}, y)$ and $\hat{L}_2(\hat{\xi}, y)$ in the above expression (3.85), we get

$$\begin{aligned}\widehat{H}(\hat{\xi}, y, \hat{L}_1, \hat{L}_2, \hat{V}, \varepsilon) &= \hat{V}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2\varepsilon), y) - \hat{V}(\hat{\xi}, y_{k-1}) + \\ &\frac{1}{2}\left\|\hat{L}_1^T(\hat{\xi}, y, \varepsilon)\hat{V}^T_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y) + (y - \tilde{h}_{21}(\hat{\xi}, \varepsilon) - h_{22}(\hat{\xi}, \varepsilon))\right\|^2 + \\ &\frac{1}{2}\left\|\frac{1}{\varepsilon}\hat{L}_2^T(\hat{\xi}, y, \varepsilon)\hat{V}^T_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y) + (y - \tilde{h}_{21}(\hat{\xi}, \varepsilon) - h_{22}(\hat{\xi}, \varepsilon))\right\|^2 - \\ &\frac{1}{2}\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y)\hat{L}_1(\hat{\xi}, y, \varepsilon)\hat{L}_1^T(\hat{\xi}, y, \varepsilon)\hat{V}^T_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y) - \\ &\frac{1}{2\varepsilon^2}\hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y)\hat{L}_2(\hat{\xi}, y, \varepsilon)\hat{L}_2^T(\hat{\xi}, y)\hat{V}^T_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y) + \\ &\frac{1}{2}\|\tilde{z}\|^2_{(W-2I)}\end{aligned}$$

Therefore, taking the filter gains as

$$\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y)\hat{L}_1^\star(\hat{\xi}, y, \varepsilon) = -(y - \tilde{h}_{21}(\hat{\xi}, \varepsilon) - h_{22}(\hat{\xi}, \varepsilon))^T, \qquad (3.86)$$

$$\hat{V}^T_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y)\hat{L}_2^\star(\hat{\xi}, y, \varepsilon) = -\varepsilon(y - \tilde{h}_{21}(\hat{\xi}, \varepsilon) - h_{22}(\hat{\xi}, \varepsilon))^T, \qquad (3.87)$$

minimizes $\widehat{H}(., ., \hat{L}_1, \hat{L}_2, ., .)$. Next, substituting the above optimal gains in (3.86) and setting

$$\widehat{H}(\hat{\xi}, y, \hat{L}_1^\star, \hat{L}_2^\star, \hat{V}, \varepsilon) = 0,$$

results in the following DHJE

$$\hat{V}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\tfrac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y) - \hat{V}(\hat{\xi},y_{k-1}) + \tfrac{1}{2}(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon))^T\times$$
$$(W-4I)(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon)) = 0, \quad \hat{V}(0,0,0)=0. \tag{3.88}$$

We then have the following result.

Proposition 3.2.2. *Consider the nonlinear discrete system (3.72) and the $\mathcal{H}_2$ filtering problem for this system. Suppose the plant $\mathbf{P}^{\mathbf{da}}_{\mathbf{sp}}$ is locally asymptotically stable about the equilibrium-point $x=0$ and zero-input observable. Further, suppose there exist a local diffeomorphism φ that transforms the system to the partially decoupled form (3.74), a C^1 positive-semidefinite function $\hat{V}:\hat{N}\times\hat{\Upsilon}\to\Re_+$ locally defined in a neighborhood $\hat{N}\times\hat{\Upsilon}\subset\mathcal{X}\times\mathcal{Y}$ of the origin $(\hat{\xi},y)=(0,0)$, and matrix functions $\hat{L}_i:\hat{N}\times\hat{\Upsilon}\to\Re^{n_i\times m}$, $i=1,2$, satisfying the DHJBE (3.88) together with the side-conditions (3.86), (3.87). Then the filter $\mathbf{F}^{da}_{1c}$ solves the $\mathcal{H}_2$ filtering problem for the system locally in $\hat{N}$.*

Proof: The optimality of the filter gains $\hat{L}_1^\star$, $\hat{L}_2^\star$ has already been shown above. It remains to prove asymptotic convergence of the estimation error vector. Accordingly, let $\hat{V}\geq 0$ be a C^2 solution of the DHJBE (3.88). Then, consider the time-variation of $\hat{V}$ along a trajectory of (3.79), with $\hat{L}_1=\hat{L}_1^\star$, $L_2=\hat{L}_2^\star$, we get

$$\begin{aligned}
\hat{V}(\hat{\xi}_{1,k+1},\hat{\xi}_{2,k+1},y) \;\approx\;& \hat{V}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y) + \\
& \hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\hat{L}_1(\hat{\xi},y,\varepsilon)(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon)) + \\
& \hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\hat{L}_2^\star(\hat{\xi},y,\varepsilon)(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon)) \\
=\;& \hat{V}(\hat{\xi}_{1,k},\hat{\xi}_{2,k},y_{k-1}) - \frac{1}{2}\|\tilde{z}_k\|^2_W,
\end{aligned}$$

where we have used the first-order Taylor approximation in the above, and the last equality follows from using the DHJBE (3.88). Moreover, the last equality also implies

$$\hat{V}(\hat{\xi}_{1,k+1},\hat{\xi}_{2,k+1},y) - \hat{V}(\hat{\xi}_{1,k},\hat{\xi}_{2,k},y_{k-1}) = -\frac{1}{2}\|\tilde{z}_k\|^2_W.$$

Therefore, the filter dynamics is stable, and $V(\hat{\xi}, y)$ is non-increasing along a trajectory of (3.79). Further, the condition that $\hat{V}(\hat{\xi}_{1,k+1}, \hat{\xi}_{2,k+1}, y) \equiv \hat{V}(\hat{\xi}_{1,k}, \hat{\xi}_{2,k}, y_{k-1})$ $\forall k \geq k_s$ implies that $\tilde{z}_k \equiv 0$, which further implies that $y_k = \tilde{h}_{21}(\hat{\xi}_{1,k}) + \tilde{h}_{22}(\hat{\xi}_{2,k})$ $\forall k \geq k_s$. By the zero-input observability of the system, this implies that $\hat{\xi} = \xi$. Finally, since φ is invertible and $\varphi(0) = 0$, $\hat{\xi} = \xi$ implies $\hat{x} = \varphi^{-1}(\hat{\xi}, \varepsilon) = \varphi^{-1}(\xi, \varepsilon) = x$. $\square$

Next, we consider the limiting behavior of the filter (3.79) and the corresponding DHJBE (3.88). Letting $\varepsilon \downarrow 0$, we obtain from (3.79),

$$0 = \tilde{f}_2(\hat{\xi}_{2,k}) + L_2(\hat{\xi}_k, y_k)(y_k - \tilde{h}_{21}(\hat{\xi}_k) - \tilde{h}_{22}(\hat{\xi}_k))$$

and since $\tilde{f}_2(.)$ is asymptotically stable, we have $\hat{\xi}_2 \to 0$. Therefore $H(.,.,.,.,.)$ in (3.82) becomes

$$\begin{aligned} H_0(\hat{\xi}, y, L_1, L_2, V, 0) \;\; = \;\; & V\Big(\tilde{f}_1(\hat{\xi}_1) + L_1(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}_1) - h_{22}(\hat{\xi}_2)), 0, y\Big) - V(\hat{\xi}, y_{k-1}) + \\ & \frac{1}{2}\|z\|^2_W. \end{aligned} \tag{3.89}$$

A first-order Taylor approximation of this Hamiltonian about $(\tilde{f}_1(\hat{\xi}_1), 0, y)$ similarly yields

$$\begin{aligned} \widehat{H}_0(\hat{\xi}, y, \hat{L}_1, \hat{L}_2, \bar{V}, 0) \;\; = \;\; & \bar{V}(\tilde{f}_1(\hat{\xi}_1), 0, y) + \bar{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1), 0, y)\hat{L}_1^T(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}) - h_{22}(\hat{\xi})) - \\ & \bar{V}(\hat{\xi}, y_{k-1}) + \frac{1}{2}\|\tilde{z}\|^2_W + O(\|\hat{\xi}\|^3). \end{aligned} \tag{3.90}$$

for some corresponding C^1-function $\bar{V} : \bar{N} \times \bar{Y} \to \Re_+$, $\bar{N} \times \bar{Y} \subset \mathcal{X} \times \mathcal{Y}$. Minimizing again this Hamiltonian, we obtain the optimal gain $\hat{L}^\star_{10}$ given by

$$\bar{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1), 0, y)\hat{L}^\star_{10}(\hat{\xi}, y) = -(y - \tilde{h}_{21}(\hat{\xi}) - h_{22}(\hat{\xi}))^T, \tag{3.91}$$

where $\bar{V}$ satisfies the reduced-order DHJBE

$$\begin{aligned} & \bar{V}(\tilde{f}_1(\hat{\xi}_1), 0, y) - \bar{V}(\hat{\xi}, y_{k-1}) + \frac{1}{2}(y - \tilde{h}_{21}(\hat{\xi}) - h_{22}(\hat{\xi}))^T(W - 2I)(y - \tilde{h}_{21}(\hat{\xi}) - h_{22}(\hat{\xi})) = 0, \\ & \bar{V}(0, 0, 0) = 0. \end{aligned} \tag{3.92}$$

The corresponding reduced-order filter is given by

$$\bar{\mathbf{F}}^{da}_{1r} : \left\{ \dot{\hat{\xi}}_1 \;=\; \tilde{f}_1(\hat{\xi}_1) + \hat{L}^{\star}_{10}(\hat{\xi}_1, y)[y - \tilde{h}_{21}(\hat{\xi}) - \tilde{h}_{22}(\hat{\xi})] + O(\varepsilon). \right. \tag{3.93}$$

Moreover, since the gain $\hat{L}^{\star}_{10}$ is such that the estimation error $e_k = y_k - \tilde{h}_{21}(\hat{\xi}_k) - \tilde{h}_{22}(\hat{\xi}_k) \to 0$, and the vector-field $\tilde{f}_2(\hat{\xi}_2)$ is locally asymptotically stable, we have $\hat{L}^{\star}_2(\hat{\xi}_k, y_k) \to 0$ as $\varepsilon \downarrow 0$. Correspondingly, the solution $\bar{V}$ of the DHJBE (3.92) can be represented as the asymptotic limit of the solution of the DHJBE (3.88) as $\varepsilon \downarrow 0$, i.e.,

$$\hat{V}(\hat{\xi}, y) = \bar{V}(\hat{\xi}_1, y) + O(\varepsilon).$$

We can also specialize the result of Proposition 3.2.2 to the following discrete-time linear singularly-perturbed system (DLSPS) (Aganovic, 1996), (Kim 2002), (Lim, 1996), (Sadjadi, 1990) in the slow coordinate:

$$\mathbf{P}^l_{dsp} \;:\; \left\{ \begin{array}{rcl} x_{1,k+1} & = & A_1 x_{1,k} + A_{12} x_{2,k} + B_{11} w_k; \quad x_1(k_0) = x^{10} \\ \varepsilon x_{2,k+1} & = & A_{21} x_{1,k} + (\varepsilon I_{n_2} + A_2) x_{2,k} + B_{21} w_k; \quad x_2(k_0) = x^{20} \\ y_k & = & C_{21} x_{1,k} + C_{22} x_{2,k} + w_k \end{array} \right. \tag{3.94}$$

where $A_1 \in \Re^{n_1 \times n_1}$, $A_{12} \in \Re^{n_1 \times n_2}$, $A_{21} \in \Re^{n_2 \times n_1}$, $A_2 \in \Re^{n_2 \times n_2}$, $B_{11} \in \Re^{n_1 \times s}$, and $B_{21} \in \Re^{n_2 \times s}$, while the other matrices have compatible dimensions. Then, an explicit form of the required transformation φ above is given by the Chang transformation (Chang, 1972):

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} I_{n_1} - \varepsilon \mathsf{HL} & -\varepsilon \mathsf{H} \\ \mathsf{L} & I_{n_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{3.95}$$

where the matrices L and H satisfy the equations

$$\begin{array}{rcl} 0 & = & (\varepsilon I_{n_2} + A_2)\mathsf{L} - A_{21} - \varepsilon \mathsf{L}(A_1 - A_{12}\mathsf{L}) \\ 0 & = & -\mathsf{H}[(\varepsilon I_{n_2} + A_2) + \varepsilon \mathsf{L} A_{12}] + A_{12} + \varepsilon (A_1 - A_{12}\mathsf{L})\mathsf{H}. \end{array}$$

The system is then represented in the new coordinates by

$$\tilde{\mathbf{P}}^{l}_{dsp}\ :\ \begin{cases} \xi_{1,k+1} &=& \tilde{A}_1\xi_{1,k} + \tilde{B}_{11}w_k; \quad \xi_1(k_0) = \xi^{10} \\ \varepsilon\xi_{2,k+1} &=& \tilde{A}_2\xi_{2,k} + \tilde{B}_{21}w_k; \quad \xi_2(k_0) = \xi^{20} \\ y_k &=& \tilde{C}_{21}x_{1,k} + \tilde{C}_{22}x_{2,k} + w_k, \end{cases} \tag{3.96}$$

where

$$\begin{aligned} \tilde{A}_1 &= A_1 - A_{12}\mathsf{L} = A_1 - A_{12}(\varepsilon I_{n_2} + A_2)^{-1}A_{21} + O(\varepsilon) \\ \tilde{B}_{11} &= B_{11} - \varepsilon\mathsf{HL}B_{11} - \mathsf{H}B_{21} = B_{11} - A_{12}A_2^{-1}B_{21} + O(\varepsilon) \\ \tilde{A}_2 &= (\varepsilon I_{n_2} + A_2) + \varepsilon\mathsf{L}A_{12} = A_2 + O(\varepsilon) \\ \tilde{B}_{21} &= B_{21} + \varepsilon\mathsf{L}B_{11} = B_{21} + O(\varepsilon) \\ \tilde{C}_{21} &= C_{21} - C_{22}\mathsf{L} = C_{21} - C_{22}(\varepsilon I_{n_2} + A_2)^{-1}A_{21} + O(\varepsilon) \\ \tilde{C}_{22} &= C_{22} + \varepsilon(C_{21} - C_{22})\mathsf{H} = C_{22} + O(\varepsilon). \end{aligned}$$

Adapting the filter (3.79) to the system (3.96) yields the following filter

$$\mathbf{F}^{dl}_{1c} : \begin{cases} \hat{\xi}_{1,k+1} &=& \tilde{A}_1\hat{\xi}_{1,k} + \hat{L}_1(y_k - \tilde{C}_{21}\hat{\xi}_{1,k} - \tilde{C}_{22}\hat{\xi}_{2,k}) \\ \varepsilon\hat{\xi}_{2,k+1} &=& \tilde{A}_2\hat{\xi}_{2,k} + \hat{L}_2(y_k - \tilde{C}_{21}\hat{\xi}_{1,k} - \tilde{C}_{22}\hat{\xi}_{2,k}). \end{cases} \tag{3.97}$$

Taking

$$\hat{V}(\hat{\xi}, y) = \frac{1}{2}(\hat{\xi}_1^T\hat{P}_1\hat{\xi}_1 + \hat{\xi}_2^T\hat{P}_2\hat{\xi}_2 + y^T\hat{Q}y),$$

for some symmetric positive-definite matrices $\hat{P}_1$, $\hat{P}_2$, $\hat{Q}$, the DHJBE (3.88). Consequently, we have the following Corollary to Proposition 3.2.2. We may assume without loss of generality that the covariance of the noise $W = I$.

Corollary 3.2.1. *Consider the DLSPS (3.94) and the $\mathcal{H}_2$ filtering problem for this system. Suppose the plant $\mathbf{P}^{l}_{sp}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and observable. Suppose further, it is transformable to the form (3.96), and there exist symmetric positive-definite matrices $\hat{P}_1 \in \Re^{n_1\times n_1}$, $\hat{P}_2 \in \Re^{n_2\times n_2}$, and $\hat{Q} \in \Re^{m\times m}$ and matrices*

$\hat{L}_1 \in \Re^{n_1 \times m}$, $\hat{L}_2 \in \Re^{n_2 \times m}$ *satisfying the LMIs*

$$\begin{bmatrix} \tilde{A}_1^T \hat{P}_1 \tilde{A}_1 - \hat{P}_1 - 3\tilde{C}_{21}^T \tilde{C}_{21} & -\tilde{C}_{22}^T \tilde{C}_{21} & 3\tilde{C}_{21}^T & 0 \\ -\tilde{C}_{21}^T \tilde{C}_{22} & \frac{1}{\varepsilon^2} \tilde{A}_2^T \hat{P}_2 \tilde{A}_2 - \hat{P}_2 - 3\tilde{C}_{22}^T \tilde{C}_{22} & 3\tilde{C}_{22}^T & 0 \\ 3\tilde{C}_{21} & 3\tilde{C}_{22} & -3I & Q \\ 0 & 0 & Q^T & 0 \end{bmatrix} \leq 0 \tag{3.98}$$

$$\begin{bmatrix} 0 & 0 & \frac{1}{2}(\tilde{A}_1^T \hat{P}_1 L_1 - \tilde{C}_{21}^T) \\ 0 & 0 & -\frac{1}{2}\tilde{C}_{22}^T \\ \frac{1}{2}(\tilde{A}_1^T \hat{P}_1 L_1 - \tilde{C}_{21}^T)^T & -\frac{1}{2}\tilde{C}_{22}^T & (1-\delta_1)I \end{bmatrix} \leq 0 \tag{3.99}$$

$$\begin{bmatrix} 0 & 0 & -\frac{1}{2}\tilde{C}_{21}^T \\ 0 & 0 & \frac{1}{2\varepsilon}(\tilde{A}_2^T \hat{P}_2 L_2 - \tilde{C}_{22}^T) \\ -\frac{1}{2}\tilde{C}_{21} & \frac{1}{2\varepsilon}(\tilde{A}_2^T \hat{P}_2 L_2 - \tilde{C}_{22}^T)^T & (1-\delta_2)I \end{bmatrix} \leq 0 \tag{3.100}$$

for some numbers $\delta_1, \delta_2 \geq 1$. *Then the filter* $\mathbf{F}_{1c}^{dl}$ *solves the* $\mathcal{H}_2$ *filtering problem for the system.*

Similarly, for the reduced-order filter (3.93) and the DHJBE (3.92), we have respectively

$$\mathbf{F}_{1r}^{dl} : \left\{ \; \hat{\xi}_{1,k+1} \;\; = \;\; \tilde{A}_1 \hat{\xi}_{1,k} + \hat{L}_{10}^{\star}(y_k - \tilde{C}_{21}\hat{\xi}_{1,k} - \tilde{C}_{22}\hat{\xi}_{2,k}), \right. \tag{3.101}$$

$$\begin{bmatrix} \tilde{A}_1^T \hat{P}_{10} \tilde{A}_1 - \hat{P}_{10} - 3\tilde{C}_{21}^T \tilde{C}_{21} & -\tilde{C}_{22}^T \tilde{C}_{21} & 3\tilde{C}_{21}^T & 0 \\ -\tilde{C}_{21}^T \tilde{C}_{22} & \frac{1}{\varepsilon^2} \tilde{A}_2^T \hat{P}_{20} \tilde{A}_2 - \hat{P}_{20} - 3\tilde{C}_{22}^T \tilde{C}_{22} & 3\tilde{C}_{22}^T & 0 \\ 3\tilde{C}_{21} & 3\tilde{C}_{22} & -3I & \hat{Q}_{10} \\ 0 & 0 & \hat{Q}_{10}^T & 0 \end{bmatrix} \leq 0 \tag{3.102}$$

$$\begin{bmatrix} 0 & 0 & \frac{1}{2}(\tilde{A}_1^T \hat{P}_{10} L_1 - \tilde{C}_{21}^T) \\ 0 & 0 & -\frac{1}{2}\tilde{C}_{22}^T \\ \frac{1}{2}(\tilde{A}_1^T \hat{P}_{10} L_1 - \tilde{C}_{21}^T)^T & -\frac{1}{2}\tilde{C}_{22}^T & (1-\delta_{10})I \end{bmatrix} \leq 0, \tag{3.103}$$

for some symmetric positive-definite matrices $\hat{P}_{10}$, $\hat{Q}_{10}$, gain matrix $\hat{L}_{10}$ and some number $\delta_{10} \geq 1$.

Proposition 3.2.2 has not yet exploited the benefit of the coordinate transformation in designing the filter (3.79) for the system (3.74). We shall now design separate reduced-order filters for the decomposed subsystems which should be more efficient than the previous one. If we let $\varepsilon \downarrow 0$ in (3.74) and obtain the following reduced system model:

$$\widetilde{\mathbf{P}}^{\mathbf{a}}_{\mathbf{r}} : \begin{cases} \xi_{1,k+1} &= \tilde{f}_1(\xi_{1,k}) + \tilde{g}_{11}(\xi_k)w_k \\ 0 &= \tilde{f}_2(\xi_{2,k}) + \tilde{g}_{21}(\xi_k)w_k \\ y &= \tilde{h}_{21}(\xi_k) + \tilde{h}_{22}(\xi_k) + \tilde{k}_{21}(\xi_k)w_k. \end{cases} \tag{3.104}$$

Then, we assume the following (Khalil, 1985), (Kokotovic, 1986).

Assumption 3.2.1. *The system (3.72), (3.104) is in the "standard form", i.e., the equation*

$$0 = \tilde{f}_2(\xi_2) + \tilde{g}_{21}(\xi)w \tag{3.105}$$

has $l \geq 1$ distinct roots, we can denote any one of these solutions by

$$\bar{\xi}_2 = q(\xi_1, w). \tag{3.106}$$

for some C^1 function $q : \mathcal{X} \times \mathcal{W} \to \mathcal{X}$.

Under Assumption 3.2.1, we obtain the reduced-order slow subsystem

$$\mathbf{P}^{\mathbf{a}}_{\mathbf{r}} : \begin{cases} \xi_{1,k+1} &= \tilde{f}_1(\xi_{1,k}) + \tilde{g}_{11}(\xi_{1,k}, q(\xi_{1,k}, w_k))w_k + O(\varepsilon) \\ y_k &= \tilde{h}_{21}(\xi_{1,k}, q(\xi_{1,k}, w_k)) + \tilde{h}_{22}(\xi_{1,k}, q(\xi_{1,k}, w_k)) + \\ & \quad \tilde{k}_{21}(\xi_{1,k}, q(\xi_{1,k}, w_k))w_k + O(\varepsilon) \end{cases} \tag{3.107}$$

and a boundary-layer (or quasi steady-state) subsystem as

$$\bar{\xi}_{2,m+1} = \tilde{f}_2(\bar{\xi}_{2,m}, \varepsilon) + \tilde{g}_{21}(\xi_{1,m}, \bar{\xi}_{2,m})w_m \tag{3.108}$$

where $m = \lfloor k/\varepsilon \rfloor$ is a stretched-time parameter. This subsystem is guaranteed to be asymptotically stable for $0 < \varepsilon < \varepsilon^\star$ (see Theorem 8.2 in Ref. (Khalil, 1985)) if the original system (3.72) is asymptotically stable.

We can then proceed to redesign the filter (3.79) for the composite system (3.107), (3.108) separately as

$$\widetilde{\mathbf{F}}^{da}_{3c} : \begin{cases} \breve{\xi}_{1,k+1} &= \tilde{f}_1(\breve{\xi}_{1,k}) + \breve{L}_1(\breve{\xi}_{1,k}, y_k)[y_k - \widetilde{h}_{21}(\breve{\xi}_{1,k}) - \widetilde{h}_{22}(\breve{\xi}_{1,k})] \\ \varepsilon\dot{\breve{\xi}}_{2,k+1} &= \tilde{f}_2(\breve{\xi}_{2,k}, \varepsilon) + \breve{L}_2(\breve{\xi}_{2,k}, y_k, \varepsilon)[y_k - \tilde{h}_{21}(\breve{\xi}_k) - \tilde{h}_{22}(\breve{\xi}_k)], \end{cases} \tag{3.109}$$

where

$$\widetilde{h}_{21}(\breve{\xi}_{1,k}) = \tilde{h}_{21}(\breve{\xi}_{1,k}, q(\breve{\xi}_{1,k}, 0)), \quad \widetilde{h}_{22}(\breve{\xi}_{1,k}) = \tilde{h}_{21}(\breve{\xi}_{1,k}, q(\breve{\xi}_{1,k}, 0)).$$

Notice also that ξ_2 cannot be estimated from (3.106) since this is a "quasi-steady-state" approximation. Then, using a similar approximation procedure as in Proposition 3.2.2 we arrive at the following result.

Theorem 3.2.1. *Consider the nonlinear system (3.72) and the $\mathcal{H}_2$ estimation problem for this system. Suppose the plant $\mathbf{P}^{\mathbf{da}}_{\mathbf{sp}}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and zero-input observable. Further, suppose there exists a local diffeomorphism φ that transforms the system to the partially decoupled form (3.74), and Assumption 3.2.1 holds. In addition, suppose there exist C^1 positive-definite functions $\breve{V}_i : \breve{N}_i \times \breve{\Upsilon}_i \to \Re_+$, $i = 1, 2$, locally defined in neighborhoods $\breve{N}_i \times \breve{\Upsilon}_i \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\breve{\xi}_i, y) = (0, 0)$, $i = 1, 2$ respectively, and matrix functions $\breve{L}_i : \breve{N}_i \times \breve{\Upsilon}_i \to \Re^{n_i \times m}$, $i = 1, 2$ satisfying the DHJBEs:*

$$\begin{aligned} &\breve{V}_1(\tilde{f}_1(\breve{\xi}_1), y) - \breve{V}_1(\breve{\xi}_1, y_{k-1}) + \frac{1}{2}(y - \widetilde{h}_{21}(\breve{\xi}_1) - \widetilde{h}_{22}(\breve{\xi}_1))^T (W - 4I)(y - \widetilde{h}_{21}(\breve{\xi}_1) - \\ &\qquad \widetilde{h}_{22}(\breve{\xi}_1)) = 0, \quad \breve{V}_1(0, 0) = 0, \end{aligned} \tag{3.110}$$

$$\begin{aligned} &\breve{V}_2(\frac{1}{\varepsilon}\tilde{f}_2(\breve{\xi}_2, \varepsilon), y) - \breve{V}_2(\breve{\xi}_2, y_{k-1}) + \frac{1}{2}(y - \tilde{h}_{21}(\breve{\xi}) - h_{22}(\breve{\xi}))^T (W - 4I)(y - \tilde{h}_{21}(\breve{\xi}) - \\ &\qquad \tilde{h}_{22}(\breve{\xi})) = 0, \quad \breve{V}_2(0, 0) = 0, \end{aligned} \tag{3.111}$$

together with the side-conditions

$$\hat{V}_{1, \hat{\xi}_1}(\tilde{f}_1(\breve{\xi}_1), y)\breve{L}^\star_1(\breve{\xi}_1, y) \;=\; -(y - \widetilde{h}_{21}(\breve{\xi}_1) - \widetilde{h}_{22}(\hat{\xi}))^T \tag{3.112}$$

$$\breve{V}_{2,\breve{\xi}_2}(\frac{1}{\varepsilon}\tilde{f}_2(\breve{\xi}_2,\varepsilon),y)\breve{L}_2^\star(\breve{\xi}_2,y,\varepsilon) \;=\; -\varepsilon(y-\tilde{h}_{21}(\breve{\xi})-\tilde{h}_{22}(\breve{\xi}))^T. \tag{3.113}$$

Then the filter $\widetilde{\mathbf{F}}^{\mathbf{da}}_{\mathbf{3c}}$ *solves the* $\mathcal{H}_2$ *filtering problem for the system locally in* $\cup\breve{N}_i$*.*

Proof: We define separately two Hamiltonian functions $H_i : \mathcal{X}\times\mathcal{Y}\times\Re^{n_i\times m}\times\Re \to \Re,\ i=1,2$ for each of the two separate components of the filter (3.109). Then, the rest of the proof follows along the same lines as Proposition 3.2.2. □

Remark 3.2.2. *Comparing (3.112), (3.110) and (3.91), (3.92), we see that the two reduced-order filters approximations are similar. Notice also that* $\breve{\xi}_1$ *appearing in (3.113), (3.110) is not considered as an additional variable, because it is assumed to be known from (3.109a), and therefore is regarded as a parameter.*

3.2.3 Discrete-time Aggregate Filters

Similarly, in the absence of the coordinate transformation, φ, discussed in the previous section, a filter has to be designed to solve the problem for the aggregate system (3.72). We discuss this class of filters in this section. Accordingly, consider the following class of filters:

$$\mathbf{F}^{da}_{3ag}\ :\ \begin{cases} \grave{x}_{1,k+1} &= f_1(\grave{x}_k)+\grave{L}_1(\grave{x}_k,y_k,\varepsilon)[y_k-h_{21}(\grave{x}_{1,k})-h_{22}(\grave{x}_{2,k})]; \\ & \quad \grave{x}_1(k_0)=\bar{x}^{10} \\ \varepsilon\grave{x}_{2,k+1} &= f_2(\grave{x}_k,\varepsilon)+\grave{L}_2(\grave{x}_k,y_k,\varepsilon)[y_k-h_{21}(\grave{x}_{1,k})-h_{22}(\grave{x}_{2,k})]; \\ & \quad \grave{x}_2(k_0)=\bar{x}^{20} \\ \grave{z}_k &= y_k-h_{21}(\grave{x}_{1,k})-h_{22}(\grave{x}_{2,k}) \end{cases} \tag{3.114}$$

where $\grave{L}_1, \grave{L}_2 \in \Re^{n\times m}$ are the filter gains, and $\grave{z}$ is the new penalty variable. We can repeat the same kind of derivation above to arrive at the following.

Theorem 3.2.2. *Consider the nonlinear system (3.72) and the* $\mathcal{H}_2$ *estimation problem for this system. Suppose the plant* $\mathbf{P}^{\mathbf{da}}_{\mathbf{sp}}$ *is locally asymptotically stable about the equilibrium-point* $x=0$*, and zero-input observable. Further, suppose there exist a* C^1 *positive-definite function* $\grave{V}:\grave{N}\times\grave{\Upsilon}\to\Re_+$*, locally defined in a neighborhood* $\grave{N}\times\grave{\Upsilon}\subset\mathcal{X}\times\mathcal{Y}$ *of the origin*

$(\grave{x}_1, \grave{x}_2, y) = (0,0,0)$, *and matrix functions* $\grave{L}_i : \grave{N} \times \grave{\Upsilon} \to \Re^{n_i \times m}$, $i = 1,2$, *satisfying the DHJBE:*

$$\grave{V}(f_1(\grave{x}), \frac{1}{\varepsilon} f_2(\grave{x},\varepsilon), y) - \grave{V}(\grave{x}, y_{k-1}) + \frac{1}{2}(y - h_{21}(\grave{x}_1) - h_{22}(\grave{x}_2))^T (W - 4I)(y - h_{22}(\grave{x}_2)) = 0, \quad \grave{V}(0,0,0) = 0, \tag{3.115}$$

together with the side-conditions

$$\grave{V}_{\grave{x}_1}(f_1(\grave{x}), \frac{1}{\varepsilon} f_2(\grave{x},\varepsilon), y)\grave{L}_1^\star(\grave{x}, y, \varepsilon) = -(y - h_{21}(\grave{x}_1) - h_{22}(\grave{x}_2))^T, \tag{3.116}$$

$$\grave{V}_{\grave{x}_2}(f_1(\grave{x}), \frac{1}{\varepsilon} f_2(\grave{x},\varepsilon), y)\grave{L}_2^\star(\grave{x}, y, \varepsilon) = -\varepsilon(y - h_{21}(\grave{x}_1) - h_{22}(\grave{x}_2))^T. \tag{3.117}$$

Then, the filter $\mathbf{F}^a_{3ag}$ *solves the* $\mathcal{H}_2$ *filtering problem for the system locally in* $\grave{N}$.

Proof: Proof follows along the same lines as Proposition 3.2.2. □

The result of Theorem 3.2.2 can similarly be specialized to the linear systems $\mathbf{P}^{dl}_{sp}$ in the following Corollary. Again we may assume without loss of generality that $W = I$.

Corollary 3.2.2. *Consider the DLSPS (3.94) and the* $\mathcal{H}_2$ *filtering problem for this system. Suppose the plant* $\mathbf{P}^l_{dsp}$ *is locally asymptotically stable about the equilibrium-point* $x = 0$ *and zero-input observable. Suppose further, there exist symmetric positive-definite matrices* $\grave{P}_1 \in \Re^{n_1 \times n_1}$, $\grave{P}_2 \in \Re^{n_2 \times n_2}$, $\grave{Q}, \grave{R} \in \Re^{m \times m}$, *and matrices* $\grave{L}_1 \in \Re^{n_1 \times m}$, $\grave{L}_2 \in \Re^{n_2 \times m}$ *satisfying the following LMIs*

$$\begin{bmatrix} A_1^T \grave{P}_1 A_1 + \frac{1}{\varepsilon^2} A_{21}^T \grave{P}_2 A_{21} - \grave{P}_1 - 3C_{21}^T C_{21} & A_1^T \grave{P}_1 A_{12} + \frac{1}{\varepsilon^2} A_{21}^T \grave{P}_2 \bar{A}_2 - 3C_{22}^T C_{21} & 3C_{21}^T & 0 \\ A_{12}^T \grave{P}_1 A_1 + \frac{1}{\varepsilon^2} \bar{A}_2^T \grave{P}_2 A_{21} - 3C_{21}^T C_{22} & A_{12}^T \grave{P}_1 A_{12} + \frac{1}{\varepsilon^2} \bar{A}_2^T \grave{P}_2 \bar{A}_2 - \grave{P}_2 - 3C_{22}^T C_{22} & 3C_{22}^T & 0 \\ 3C_{21} & 3C_{22} & -3I & Q \\ 0 & 0 & Q^T & 0 \end{bmatrix} \leq 0 \tag{3.118}$$

$$\begin{bmatrix} 0 & 0 & \frac{1}{2}(A_1^T\grave{P}_1\grave{L}_1 - C_{21}^T) \\ 0 & 0 & -\frac{1}{2}C_{22}^T \\ \frac{1}{2}(A_1^T\grave{P}_1\grave{L}_1 - C_{21}^T)^T & -\frac{1}{2}C_{22}^T & (1-\mu_1)I \end{bmatrix} \leq 0 \tag{3.119}$$

$$\begin{bmatrix} 0 & 0 & -\frac{1}{2}C_{21}^T \\ 0 & 0 & \frac{1}{2\varepsilon}(\bar{A}_2^T\grave{P}_2\grave{L}_2 - C_{22}^T) \\ -\frac{1}{2}C_{21} & \frac{1}{2\varepsilon}(\bar{A}_2^T\grave{P}_2\grave{L}_2 - C_{22}^T)^T & (1-\mu_2)I \end{bmatrix} \leq 0. \tag{3.120}$$

Then the filter

$$\mathbf{F}_{1ag}^{dl} : \begin{cases} \grave{x}_{1,k+1} &= A_1\grave{x}_{1,k} + A_{12}\grave{x}_{2,k} + \grave{L}_1(y_k - C_{21}\grave{x}_{1,k} - C_{22}\grave{x}_{2,k}) \\ \varepsilon\grave{x}_{2,k+1} &= A_{21}\grave{x}_{1,k} + \bar{A}_2\grave{x}_{2,k} + \grave{L}_2(y_k - C_{21}\grave{x}_{1,k} - C_{22}\grave{x}_{2,k}), \end{cases} \tag{3.121}$$

where $\bar{A}_2 = (\varepsilon I_{n_2} + A_2)$ *solves the* $\mathcal{H}_2$ *filtering problem for the system.*

We also have the limiting behavior of the filter $\mathbf{F}_{3ag}^{da}$ as $\varepsilon \downarrow 0$

$$\bar{\mathbf{F}}_{5ag}^{da} : \begin{cases} \grave{x}_{1,k+1} &= f_1(\grave{x}_k) + \grave{L}_1(\grave{x}_k, y_k)[y_k - h_{21}(\grave{x}_{1,k})]; \quad \grave{x}_1(k_0) = \bar{x}^{10} \\ \grave{x}_{2,k} &\to 0, \end{cases} \tag{3.122}$$

and the DHJBE (3.115) reduces to the DHJBE

$$\grave{V}(f_1(\grave{x}), y) - \grave{V}(\grave{x}, y_{k-1}) + \frac{1}{2}(y - h_{21}(\grave{x}_1))^T(W - 2I)(y - h_{21}(\grave{x}_1)) = 0, \quad \grave{V}(0,0) = 0, \tag{3.123}$$

together with the side-conditions

$$\grave{V}_{\grave{x}_1}(f_1(\grave{x}), y)\grave{L}_1^\star(\grave{x}, y) = -(y - h_{21}(\grave{x}_1) - h_{22}(\grave{x}_2))^T \tag{3.124}$$

$$\grave{L}_2(\grave{x}, y) \to 0. \tag{3.125}$$

3.2.4 Discrete-time Push-Pull Configuration

Finally, in this subsection, we present similarly a *"push-pull"* configuration for the discrete aggregate filter presented in the above section. Since the dynamics of the second subsystem is fast, we can afford to reduce the gain of the filter for this subsystem to avoid instability, while for the slow subsystem, we can afford to increase the gain. Therefore, we consider the following filter configuration

$$\mathbf{F}^{da}_{7ag} : \begin{cases} x^\flat_{1,k+1} &= f_1(x^\flat_k) + (L^\flat_1 + L^\flat_2)(x^\flat_k, y_k, \varepsilon)(y_k - h_{21}(x^\flat_{1,k}) - h_{22}(x^\flat_{2,k})]; \\ & \quad x^\flat_1(k_0) = \bar{x}_{10} \\ \varepsilon x^\flat_{2,k+1} &= f_2(x^\flat_k, \varepsilon) + (L^\flat_1 - L^\flat_2)(x^\flat_k, y_k, \varepsilon)[y_k - h_{21}(x^\flat_{1,k}) - \\ & \quad h_{22}(x^\flat_{2,k})]; \quad x^\flat_2(k_0) = \bar{x}_{20} \\ z^\flat &= y_k - h_{21}(x^\flat_{1,k}) - h_{22}(x^\flat_{2,k}), \end{cases} \tag{3.126}$$

where $x^\flat \in \mathcal{X}$ is the filter state, $L^\flat_1 \in \Re^{n_1 \times m}$, $L^\flat_2 \in \Re^{n_2 \times m}$ are the filter gains, while all the other variables have their corresponding previous meanings and dimensions.

Consequently, going through similar manipulations as in Proposition 3.2.2 we can give a corresponding result to Theorem 3.2.2 for the push-pull configuration.

Proposition 3.2.3. *Consider the nonlinear system (3.72) and the $\mathcal{H}_2$ estimation problem for this system. Suppose the plant $\mathbf{P}^{\mathbf{da}}_{\mathbf{sp}}$ is locally asymptotically stable about the equilibrium-point $x = 0$, and zero-input observable. Further, suppose there exist a C^1 positive-semidefinite function $V^\flat : N^\flat \times \Upsilon^\flat \to \Re_+$, locally defined in a neighborhood $N^\flat \times \Upsilon^\flat \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(x^\flat_1, x^\flat_2, y) = (0, 0, 0)$, and matrix functions $L^\flat_1 \in \Re^{n_1 \times m}$, $L^\flat_2 :\in \Re^{n_2 \times m}$, satisfying the DHJBE (3.115) together with the side-conditions*

$$[V^\flat_{x^\flat_1}(f_1(x^\flat), \frac{1}{\varepsilon} f_2(x^\flat, \varepsilon)) + \varepsilon V^\flat_{x^\flat_2}(f_1(x^\flat), \frac{1}{\varepsilon} f_2(x^\flat, \varepsilon))] L^\flat_1(x^\flat, y, \varepsilon) = \\ -(y - h_{21}(x^\flat_1) - h_{22}(x^\flat_2))^T \tag{3.127}$$

$$[V^\flat_{x^\flat_1}(\tilde{f}_1(x^\flat), \frac{1}{\varepsilon} f_2(x^\flat, \varepsilon)) - \varepsilon V^\flat_{x^\flat_2}(f_1(x^\flat), \frac{1}{\varepsilon} f_2(x^\flat, \varepsilon))] L^\flat_2(x^\flat, y, \varepsilon) = \\ -(y - h_{21}(x^\flat_1) - h_{22}(x^\flat_2))^T. \tag{3.128}$$

Then, the filter $\mathbf{F}^a_{7ag}$ solves the $\mathcal{H}_2$ filtering problem for the system locally in $N^\flat$.

Remark 3.2.3. *If the nonlinear system (3.72) is in the standard form, i.e., the equivalent of*

Assumption 3.2.1 is satisfied, and there exists at least one root $\bar{x}_2 = \sigma(x_1, w)$ *to the equation*

$$0 = f_2(x_1, x_2) + g_{21}(x_1, x_2)w,$$

then reduced-order filters can also be constructed for the system similar to the result of Subection 3.2.2 and Theorem 3.2.1. Such filters would take the following form

$$\mathbf{F}^{da}_{8agr} : \begin{cases} \breve{x}_{1,k+1} &= f_1(\breve{x}_{1,k}, \sigma(\breve{x}_1, 0)) + \check{L}_1(\breve{x}_{1,k}, y_k)[y_k - h_{21}(\breve{x}_{1,k}) - \\ & \quad h_{22}(\sigma(\breve{x}_1, 0))]; \quad \breve{x}_1(k_0) = \bar{x}_{10} \\ \varepsilon\breve{x}_{2,k+1} &= f_2(\breve{x}_k, \varepsilon) + \check{L}_2(\breve{x}_k, y_k, \varepsilon)[y_k - h_{21}(\breve{x}_{1,k}) - h_{22}(\breve{x}_{2,k})]; \\ & \quad \breve{x}_2(k_0) = \bar{x}_{20} \\ \breve{z}_k &= y_k - h_{21}(\breve{x}_{1,k}) - h_{22}(\breve{x}_{2,k}). \end{cases}$$

However, this filter would fall into the class of decomposition filters, rather than aggregate, and because of this, we shall not discuss it further in this section.

In the next subsection, we consider an examples.

3.3 Examples

We consider some simple examples in this section, because of the difficulty of solving the HJBE.

Example 3.3.1. *Consider the following singularly-perturbed nonlinear system*

$$\begin{aligned} x_{1,k+1} &= x_{1,k}^{\frac{1}{3}} + x_2^{\frac{1}{5}} + x_{1,k}w_{0,k} \\ \varepsilon x_{2,k+1} &= -x_{2,k}^{\frac{2}{3}} - x_{2,k}^{\frac{2}{5}} + sin(x_{2,k})w_{0,k} \\ y_k &= x_{1,k} + x_{2,k} + w_0. \end{aligned}$$

where w_0 *is a zero-mean Gaussian white noise process with covariance* $W = I$, $\varepsilon > 0$. *We construct the aggregate filter* $\mathbf{F}^a_{3ag}$ *presented in the previous section for the above system. It can be checked that the system is locally observable, and the function* $\breve{V}(\breve{x}) = \frac{1}{2}(\breve{x}_1^2 + \varepsilon\breve{x}_2^2)$,

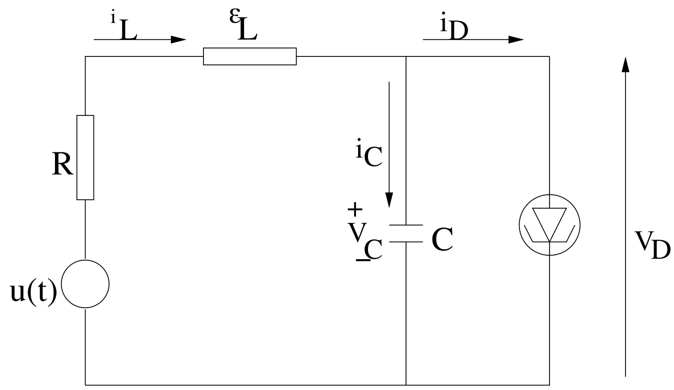

Figure 3.3 Tunnel-diode circuit.

solves the inequality form of the DHJBE (3.115) corresponding to system. Subsequently, we calculate the gains of the filter as

$$\breve{L}_1(\breve{x}, y) = -\frac{(y - \breve{x}_1 - \breve{x}_2)}{\breve{x}_1^{\frac{1}{3}} + \breve{x}_2^{\frac{1}{5}}}, \quad \breve{L}_2(\breve{x}, y) = -\frac{\varepsilon(y - \breve{x}_1 - \breve{x}_2)}{\breve{x}_1^{\frac{1}{3}} + \breve{x}_2^{\frac{1}{5}}}, \tag{3.129}$$

where the gains $\breve{L}_1$, $\breve{L}_2$ *are set equal to zero if* $|\breve{x}_1^{\frac{1}{3}} + \breve{x}_2^{\frac{1}{5}}| < \epsilon$ *(small) to avoid a singularity.*

Next, we consider the tunnel-diode example considered in (Assawinchaichote, 2004b), (Hong, 2008).

Example 3.3.2. *Consider the tunnel diode circuit example in (Assawinchaichote, 2004b), (Hong, 2008) and shown in Fig. 3.3. Suppose also the doping of the diode is such that the diode current in the circuit is given by*

$$i_D(t) = 0.01 v_D(t) + 0.05 v_D^{1/3}(t)$$

with the parasitic inductance defined by ϵ_L *and the state variables* $x_1(t) = v_C(t)$, $x_2(t) = i_L(t)$, *we have*

$$\begin{cases} C\dot{x}_1(t) &= -0.01x_1(t) - 0.05x_1^{1/3}(t) + x_2(t) + 0.1w(t) \\ \epsilon_L \dot{x}_2(t) &= -x_1(t) - Rx_2(t) \\ y(t) &= Sx(t) + w(t). \end{cases}$$

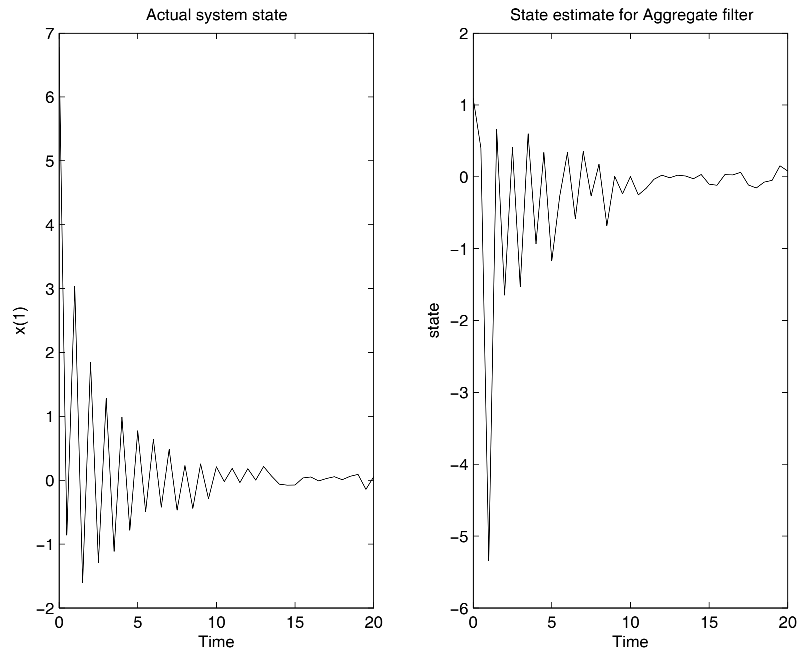

Figure 3.4 Actual state and state estimate for Reduced aggregate $\mathcal{H}_2$ filter with unknown initial condition and noise variance 0.1.

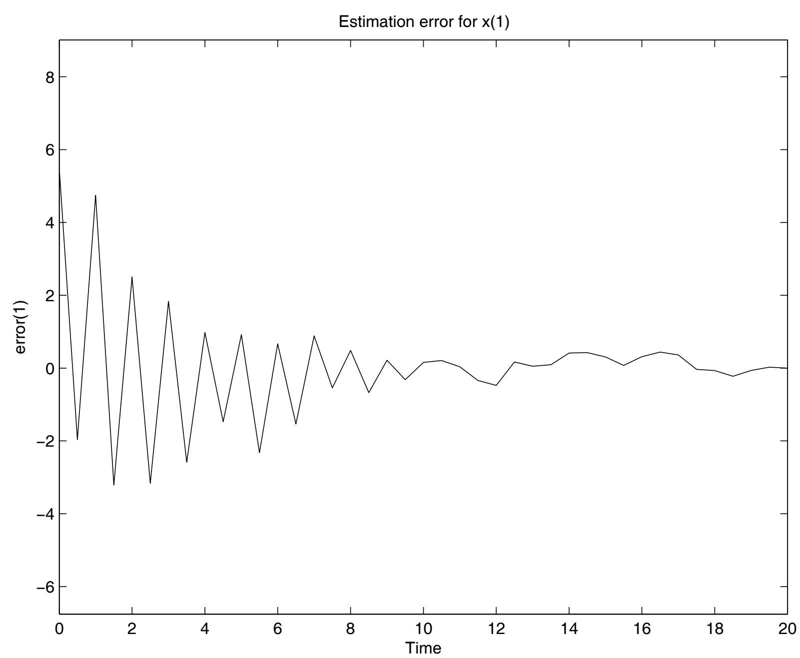

Figure 3.5 Estimation error for Reduced aggregate $\mathcal{H}_2$ filter with unknown initial condition and noise variance 0.1.

If we choose $C = 1F$ *and* $R = 0.4\Omega$, $\epsilon_L = \epsilon H = 0.01H$, $S = [1 \;\; 1]$, *then the discrete-time approximation of the above cicuit with sampling width 1 is given by*

$$\left\{ \begin{array}{rcl} x_{1,k+1} & = & x_{1,k} - 0.05x_{1,k}^{1/3} + x_{2,k} + w_k \\ \epsilon_L x_{2,k+1} & = & -x_{1,k} - 0.4x_{2,k} \\ y_k & = & x_{1,k} + x_{2,k} + w_k. \end{array} \right. \tag{3.130}$$

Suppose also that we are only interested in estimating the output voltage across the diode and capacitor $v_D(t) = x_1(t)$. *Then, we can consider the reduced-order filter* $\bar{\mathbf{F}}_{5ag}^{da}$ *for the system defined by*

$$\left\{ \begin{array}{rcl} \grave{x}_{1,k+1} & = & -1.5x_{1,k} - 0.05x_{1,k}^{1/3} + \\ & & \grave{l}_1(\grave{x}_k, y_k)(y_k + 1.5(\grave{x}_{1,k})); \\ & & \grave{x}_1(k_0) = \bar{x}^{10} \end{array} \right.$$

We can take as a local approximate solution to the DHJBE (3.115) for the above system as $\grave{V}(\grave{x}) = \frac{1}{2}(\grave{x}_1 + \varepsilon_L x_2)^2$. *Using this solution, we calculate the filter gain from (3.125) as*

$$\grave{l}_1(\grave{x}) = \frac{-(y + 1.5\grave{x}_1)}{-1.5x_1 - 0.05\grave{x}_1^{1/3}}$$

Notice that in this example, it is easier to find an approximate solution of the DHJBE (3.115) than of (3.123).

The results of the simulation with the above filter are shown on Figs. 3.4,3.5. The noise signals w_1 *are assumed to be zero mean Guassian white noises with variances* 0.1. *The initial condition for the system is also assumed to be unknown in the simulations. The results of the simulation are reasonably good, considering the fact that we are using only an approximate solution to the DHJBE.*

3.4 Conclusion

In this Chapter, we have presented a solution to the $\mathcal{H}_2$ filtering problem for affine nonlinear singularly-perturbed systems. Three types of filters have been presented, and sufficient conditions for the solvability of the problem using each filter have been given in terms of HJBEs. Both continuous-time and the discrete-time systems have been considered, and the results have also been specialized to linear systems, in which case the conditions reduce to a system of LMIs which are computationally efficient to solve. Examples have also been presented to illustrate the approach.

However, efforts would still have to be made in finding an explicit form for the coordinate transformation discussed in Subsections 2.2, and 3.2, and also in developing computationally efficient algorithms for solving the HJIEs.

CHAPTER 4

$\mathcal{H}_\infty$ FILTERING FOR SINGULARLY-PERTURBED NONLINEAR SYSTEMS

In this chapter, we discuss the $\mathcal{H}_\infty$ filtering problem for affine nonlinear singularly-perturbed systems. The $\mathcal{H}_2$ techniques discussed in the previous chapter suffer from the lack of robustness towards $\mathcal{L}_2$ bounded disturbances and other types of noise that are nonGaussian. On the other hand, the $\mathcal{H}_\infty$ filter is the optimal worst-case filter for all $\mathcal{L}_2$-bounded noise signals and is also robust against unmodelled system dynamics or uncertainties. Furthermore, $\mathcal{H}_\infty$ filtering techniques have been applied to linear and nonlinear singularly-perturbed systems by some authors (Assawinchaichote, 2004a), (Assawinchaichote, 2004b), (Lim, 1996), (Yang, 2008). In particular the references (Assawinchaichote, 2004a), (Assawinchaichote, 2004b), (Yang, 2008) have considered fuzzy T-S nonlinear singularly-perturbed systems and have used linear-matrix-inequalities (LMIs) for the filter design, which make the approach computationally very attractive. However, to the best of our knowledge, the general affine nonlinear problem has not been considered by any authors. Therefore, we propose to discss this problem in this chapter. Three types of filters will similarly be considered, and sufficient conditions for the solvability of the problem in terms of Hamilton-Jacobi-Isaacs equations (HJIEs) will be presented. An upper bound $\varepsilon^\star$ on the singular parameter ε for the stability of the nonlinear filters is also obtained using the local linearization of the nonlinear models.

In addition, there has been alot of progress in the application of nonlinear $\mathcal{H}_\infty$ techniques in control and filtering as efficient computational algorithms for solving HJIEs are being developed (Aliyu, 2003)-(Abukhalaf, 2006), (Feng, 2009), (Huang, 1999), (Sakamoto, 2008). The advantages of using the nonlinear $\mathcal{H}_\infty$ approach is that, the full nonlinear system model is utilized in determining a solution to the problem, and solutions obtained are optimal (or suboptimal) over the domain of validity of the solution to the HJIE. Hence they are more reliable, plus the additional benefit of robustness to modeling errors and disturbances. Moreover, by specializing the results developed to linear systems, we get a local approximate solution to the problem corresponding to the linearization of the system around an operating point. Indeed, using this local linarization, an upper bound $\varepsilon^\star$ on the singular parameter ε for the stability of the nonlinear filters can also be obtained. This in itself is an added motivation for considering the nonlinear techniques. But the problem also deserves consideration in its own right. Both the continuous-time and the discrete-time problems will be discussed. The Chapter is organized as follows.

4.1 $\mathcal{H}_\infty$ Filtering for Continuous-time Systems

In this Section we discuss the $\mathcal{H}_\infty$ filtering problem for continuous-time singularly-perturbed affine nonlinear systems, and in the next section, we discuss the corresponding discrete-time results. Under each section, we discuss decomposition, reduced and aggregate filters

4.1.1 Problem Definition and Preliminaries

The general set-up for studying $\mathcal{H}_\infty$ filtering problems is shown in Fig. 4.1, where $\mathbf{P}$ is the plant, while $\mathbf{F}$ is the filter. The noise signal $w \in \mathcal{P}$ is in general a bounded power signal (or $\mathcal{L}_2$ signal) which belongs to the set $\mathcal{P}$ of bounded power signals, and similarly the output $\tilde{z} \in \mathcal{P}$ is a bounded power signal. Thus, the induced norm from w to $\tilde{z}$ (the penalty variable to be defined later) is the $\mathcal{L}_\infty$-norm of the interconnected system $\mathbf{F} \circ \mathbf{P}$, i.e.,

$$\|\mathbf{F} \circ \mathbf{P}\|_{\mathcal{L}_\infty} \triangleq \sup_{0 \neq w \in \mathcal{S}} \frac{\|\tilde{z}\|_{\mathcal{P}}}{\|w\|_{\mathcal{P}}}, \tag{4.1}$$

and is defined as the $\mathcal{H}_\infty$-norm of the system for stable plant-filter pair $\mathbf{F} \circ \mathbf{P}$, where

$$\begin{aligned}
\mathcal{P} &\triangleq \{w(t) :\ w \in \mathcal{L}_\infty,\ R_{ww}(\tau),\ S_{ww}(j\omega) \text{ exist for all } \tau \text{ and all } \omega \text{ resp.},\ \|w\|_{\mathcal{P}} < \infty\}, \\
\|z\|^2_{\mathcal{P}} &\triangleq \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \|z(t)\|^2 dt.
\end{aligned}$$

At the outset, we consider the following affine nonlinear causal state-space model of the plant

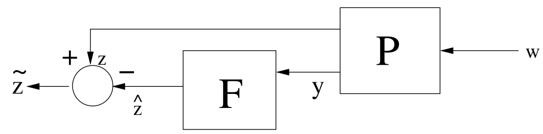

Figure 4.1 Set-up for $\mathcal{H}_\infty$ Filtering

which is defined on a manifold $\mathcal{X} \subseteq \Re^{n_1+n_2}$ with zero control input:

$$\mathbf{P}^a_{sp} : \begin{cases} \dot{x}_1 &= f_1(x_1,x_2) + g_{11}(x_1,x_2)w; \quad x_1(t_0,\varepsilon) = x_{10} \\ \varepsilon\dot{x}_2 &= f_2(x_1,x_2) + g_{21}(x_1,x_2)w; \quad x_2(t_0,\varepsilon) = x_{20} \\ y &= h_{21}(x_1) + h_{22}(x_2) + k_{21}(x_1,x_2)w, \end{cases} \tag{4.2}$$

where $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathcal{X}$ is the state vector with x_1 the slow state which is n_1-dimensional and x_2 the fast, which is n_2-dimensional; $w \in \mathcal{W} \subseteq \mathcal{L}_2([t_0,\infty),\Re^r)$ is an unknown disturbance (or noise) signal, which belongs to the set $\mathcal{W}$ of admissible exogenous inputs; $y \in \mathcal{Y} \subset \Re^m$ is the measured output (or observation) of the system, and belongs to $\mathcal{Y}$, the set of admissible measured-outputs; while $\varepsilon > 0$ is a small perturbation parameter.

The functions $\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} : \mathcal{X} \to T\mathcal{X} \subseteq \Re^{2(n_1+n_2)}$ [1], $g_{11} : \mathcal{X} \to \mathcal{M}^{n_1\times r}(\mathcal{X})$, $g_{21} : \mathcal{X} \to \mathcal{M}^{n_2\times r}(\mathcal{X})$, where $\mathcal{M}^{i\times j}$ is the ring of $i\times j$ smooth matrices over $\mathcal{X}$, $h_{21}, h_{22} : \mathcal{X} \to \Re^m$, and $k_{21} : \mathcal{X} \to \mathcal{M}^{m\times r}(\mathcal{X})$ are real C^∞ functions of x. Furthermore, we assume without any loss of generality that the system (4.2) has an isolated equilibrium-point at $(x_1,x_2) = (0,0)$ and such that $f_1(0,0) = 0$, $f_2(0,0) = 0$, $h_{21}(0,0) = h_{22}(0,0) = 0$. We also assume that there exists a unique solution $x(t,t_0,x_0,w,\varepsilon)$ $\forall t \in \Re$ for the system for all initial conditions x_0, for all $w \in \mathcal{W}$, and all $\varepsilon \in \Re$.

Moreover, to guarantee local asymptotic stability of the system (4.2) with $w = 0$, we assume that (4.2) satisfies the conditions of Theorem 8.2, (Khalil, 1985), i.e., there exists an $\varepsilon^\star > 0$ such that (4.2) is locally asymptotic stable about $x = 0$ for all $\varepsilon \in [0,\varepsilon^\star)$.

[1] For a manifold M, TM and $T^\star M$ are the tangent and cotangent bundles of M.

The $\mathcal{H}_\infty$-suboptimal local filtering/state estimation problem is defined as follows.

Definition 4.1.1. *($\mathcal{H}_\infty$-Suboptimal Local Filtering/State Estimation Problem). Find a filter,* $\mathbf{F}$*, for estimating the state $x(t)$ or a function of it, $z = h_1(x)$, from observations $\mathsf{Y}_t \triangleq \{y(\tau) : \tau \leq t\}$ of $y(\tau)$ up to time t, to obtain the estimate*

$$\hat{x}(t) = \mathbf{F}(\mathsf{Y}_t),$$

such that, the $\mathcal{H}_\infty$-norm from the input w to some suitable penalty function $\tilde{z}$ is rendered less or equal to a given number $\gamma > 0$, i.e.,

$$\int_{t_0}^{\infty} \|\tilde{z}(\tau)\|^2 dt \leq \gamma^2 \int_{t_0}^{\infty} \|w(\tau)\|^2 dt, \quad \forall w \in \mathcal{W}, \tag{4.3}$$

for all initial conditions $x_0 \in \mathcal{O} \subset \mathcal{X}$. Moreover, if the filter solves the problem for all $x_0 \in \mathcal{X}$, we say the problem is solved globally.

We shall adopt the following notion of local observability.

Definition 4.1.2. *For the nonlinear system $\mathbf{P}^a_{sp}$, we say that it is locally zero-input observable, if for all states $x_1,\ x_2 \in U \subset \mathcal{X}$ and input $w(.) = 0$*

$$y(t; x_1, w) \equiv y(t; x_2, w) \Longrightarrow x_1 = x_2,$$

where $y(., x_i, w), i = 1, 2$ is the output of the system with the initial condition $x(t_0) = x_i$. Moreover, the system is said to be zero-input observable if it is locally zero-input observable at each $x_0 \in \mathcal{X}$ or $U = \mathcal{X}$.

4.1.2 Solution to the $\mathcal{H}_\infty$ Filtering Problem Using Decomposition Filters

In this section, we present a decompostion approach to the $\mathcal{H}_\infty$ state estimation problem. We construct two time-scale filters corresponding to the decomposition of the system into a *"fast"* and *"slow"* subsystems. As in the linear case (Chang, 1972), (Gajic, 1994), (Haddad, 1976), we assume that, there exists locally a smooth invertible coordinate transformation (a diffeomorphism), $\varphi : x \mapsto \xi$, i.e.,

$$\xi_1 = \varphi_1(x), \quad \varphi_1(0) = 0, \quad \xi_2 = \varphi_2(x), \quad \varphi_2(0) = 0, \;\; \xi_1 \in \Re^{n_1}, \xi_2 \in \Re^{n_2}, \tag{4.4}$$

such that the system (4.2) can be decomposed in the form

$$\tilde{\mathbf{P}}^{\mathbf{a}}_{sp} : \begin{cases} \dot{\xi}_1 &= \tilde{f}_1(\xi_1) + \tilde{g}_{11}(\xi)w, \quad \xi_1(t_0) = \varphi_1(x_0) \\ \varepsilon\dot{\xi}_2 &= \tilde{f}_2(\xi_2) + \tilde{g}_{21}(\xi)w; \quad \xi_2(t_0) = \varphi_2(x_0) \\ y &= \tilde{h}_{21}(\xi_1) + \tilde{h}_{22}(\xi_2) + \tilde{k}_{21}(\xi)w. \end{cases} \tag{4.5}$$

Necessary conditions that such a transformation has to satisfy are given in (Aliyu, 2011c) and in Chapter 3. Then, we can proceed to design the filter based on this transformed model (4.5) with the systems states partially decoupled. Accordingly, we propose the following composite "certainty-equivalent" filter

$$\mathbf{F}^{a}_{1c} : \begin{cases} \dot{\hat{\xi}}_1 &= \tilde{f}_1(\hat{\xi}_1) + \tilde{g}_{11}(\hat{\xi})\hat{w}^\star + \hat{L}_1(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2)); \;\; \hat{\xi}_1(t_0) = \varphi_1(0) \\ \varepsilon\dot{\hat{\xi}}_2 &= \tilde{f}_2(\hat{\xi}_2) + \tilde{g}_{21}(\hat{\xi})\hat{w}^\star + \hat{L}_2(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2)) \;\; \hat{\xi}_2(t_0) = \varphi_2(0) \end{cases} \tag{4.6}$$

where $\hat{\xi} \in \mathcal{X}$ is the filter state, $\hat{w}^\star$ is the certainty-equivalent worst-case noise, $\hat{L}_1 \in \Re^{n_1 \times m}$, $\hat{L}_2 \in \Re^{n_2 \times m}$ are the filter gains, while all the other variables have their corresponding previous meanings and dimensions. We can then define the penalty variable or estimation error as

$$z = y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2). \tag{4.7}$$

The problem can then be formulated as a dynamic optimization problem (or zero-sum differential game) with the following cost functional (Basar, 1995):

$$\begin{aligned} \min_{L_1 \in \Re^{n_1 \times m}, L_2 \in \Re^{n_2 \times m}} \sup\nolimits_{w \in \mathcal{W}} \hat{J}_1(\hat{L}_1, \hat{L}_2, w) &= \frac{1}{2}\int_{t_0}^{\infty} (\|z\|^2 - \gamma^2\|w\|^2)dt, \quad s.t. \quad (4.6), \\ &\quad \text{and with } w = 0, \;\; \lim_{t\to\infty}\{\hat{\xi}(t) - \xi(t)\} = 0. \end{aligned} \tag{4.8}$$

A *"saddle-point solution"* (Basar, 1995) to the above game is said to exist, if we can find a pair of strategies $([L_1^\star, L_2^\star], w^\star)$ such that the following conditions are satisfied

$$\hat{J}_1([\hat{L}_1^\star, \hat{L}_2^\star], w) \le \hat{J}_1([\hat{L}_1^\star, \hat{L}_2^\star], w^\star) \le \hat{J}_1([\hat{L}_1, \hat{L}_2], w^\star), \ \forall \hat{L}_1 \in \Re^{n_1\times m}, \hat{L}_2 \in \Re^{n_2\times m}, \ \forall w \in \mathcal{W} \tag{4.9}$$

To solve the above problem, we form the Hamiltonian function $\hat{H} : T^\star\mathcal{X} \times T^\star\mathcal{Y} \times \mathcal{W} \times \Re^{n_1\times m} \times \Re^{n_2\times m} \to \Re$:

$$\begin{aligned}\hat{H}(\hat{\xi}, y, w, \hat{L}_1, \hat{L}_2, \hat{V}_{\hat{\xi}}^T, \hat{V}_y^T) \;\; = \;\; & \hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)[\tilde{f}_1(\hat{\xi}_1) + \tilde{g}_{11}(\hat{\xi})w + \hat{L}_1(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}_1) - h_{22}(\hat{\xi}_2))] + \\ & \frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)[\tilde{f}_2(\hat{\xi}_2) + \tilde{g}_{21}(\xi)w + \hat{L}_2(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}_1) - h_{22}(\hat{\xi}_2))] + \\ & \hat{V}_y(\hat{\xi}, y)\dot{y} + \frac{1}{2}(\|z\|^2 - \gamma^2\|w\|^2)\end{aligned} \tag{4.10}$$

for some C^1 function $\hat{V} : \mathcal{X} \times \mathcal{Y} \to \Re$. Then, applying the necessary condition for the worst-case noise, we have

$$\left.\frac{\partial \hat{H}}{\partial w}\right|_{w=\hat{w}^\star} = 0 \Longrightarrow \hat{w}^\star = \frac{1}{\gamma^2}[\tilde{g}_{11}^T(\hat{\xi})\hat{V}_{\hat{\xi}_1}^T(\hat{\xi}, y) + \frac{1}{\varepsilon}\tilde{g}_{21}^T(\hat{\xi})\hat{V}_{\hat{\xi}_2}^T(\hat{\xi}, y)]. \tag{4.11}$$

Moreover,

$$\frac{\partial^2 \hat{H}}{\partial w^2} = -\gamma^2 I \Longrightarrow \hat{H}(\hat{\xi}, y, w, \hat{L}_1, \hat{L}_2, \hat{V}_{\hat{\xi}}^T, \hat{V}_y^T) \le \hat{H}(\hat{\xi}, y, \hat{w}^\star, \hat{L}_1, \hat{L}_2, \hat{V}_{\xi}^T, \hat{V}_y^T) \ \ \forall w \in \mathcal{W}.$$

Substituting now $\hat{w}^\star$ in (4.10) and completing the squares for $\hat{L}_1$ and $\hat{L}_2$, we have

$$\begin{aligned}\hat{H}(\hat{\xi}, y, \hat{w}^\star, \hat{L}_1, \hat{L}_2, \hat{V}_{\hat{\xi}}^T, \hat{V}_y^T) = \; & \hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)\tilde{f}_1(\hat{\xi}_1) + \frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)\tilde{f}_2(\hat{\xi}_2) + \hat{V}_y(\hat{\xi}, y)\dot{y} + \\ & \frac{1}{2\gamma^2}[\hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)\tilde{g}_{11}(\hat{\xi})\tilde{g}_{11}^T(\hat{\xi})\hat{V}_{\hat{\xi}_1}^T(\hat{\xi}, y) + \frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)\tilde{g}_{11}(\xi)\tilde{g}_{21}^T(\hat{\xi})\hat{V}_{\hat{\xi}_1}^T(\hat{\xi}, y) + \\ & \frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)\tilde{g}_{21}(\xi)\tilde{g}_{11}^T(\hat{\xi})\hat{V}_{\hat{\xi}_1}^T(\hat{\xi}, y) + \frac{1}{\varepsilon^2}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)\tilde{g}_{21}(\hat{\xi})\tilde{g}_{21}^T(\hat{\xi})\hat{V}_{\hat{\xi}_2}^T(\hat{\xi}, y)] + \\ & \frac{1}{2}\|\hat{L}_1^T(\hat{\xi}, y)\hat{V}_{\hat{\xi}_1}^T(\hat{\xi}, y) + (y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))\|^2 - \\ & \frac{1}{2}\|(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))\|^2 - \frac{1}{2}\hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)\hat{L}_1(\hat{\xi}, y)\hat{L}_1^T(\hat{\xi}, y)V_{\hat{\xi}_1}^T(\hat{\xi}, y) +\end{aligned}$$

$$\frac{1}{2}\left\|\frac{1}{\varepsilon}\hat{L}_2^T(\hat{\xi},y)\hat{V}_{\hat{\xi}_2}^T(\hat{\xi},y)+(y-\tilde{h}_{21}(\hat{\xi}_1)-\tilde{h}_{22}(\hat{\xi}_2))\right\|^2+\frac{1}{2}\|z\|^2-$$
$$\frac{1}{2}\|(y-\tilde{h}_{21}(\hat{\xi}_1)-\tilde{h}_{22}(\hat{\xi}_2))\|^2-\frac{1}{2\varepsilon^2}\hat{V}_{\hat{\xi}_2}(\hat{\xi},y)\hat{L}_2(\hat{\xi},y)\hat{L}_2^T(\hat{\xi},y)\hat{V}_{\hat{\xi}_2}^T(\hat{\xi},y).$$

Thus, setting the optimal gains $\hat{L}_1^\star(\hat{\xi},y)$, $\hat{L}_2^\star(\hat{\xi},y)$ as

$$\begin{aligned}\hat{V}_{\hat{\xi}_1}(\hat{\xi},y)\hat{L}_1^\star(\hat{\xi},y) &= -(y-\tilde{h}_{21}(\hat{\xi}_1)-\tilde{h}_{22}(\hat{\xi}_2))^T, &(4.12)\\ \hat{V}_{\hat{\xi}_2}(\hat{\xi},y)\hat{L}_2^\star(\hat{\xi},y) &= -\varepsilon(y-\tilde{h}_{21}(\hat{\xi}_1)-\tilde{h}_{22}(\hat{\xi}_2))^T, &(4.13)\end{aligned}$$

minimizes the Hamiltonian (4.10) and implies that the saddle-point condition

$$\hat{H}(\hat{\xi},y,\hat{w}^\star,\hat{L}_1^\star,\hat{L}_2^\star,\hat{V}_{\hat{\xi}}^T,\hat{V}_y^T)\leq\hat{H}(\hat{\xi},y,\hat{w}^\star,\hat{L}_1,\hat{L}_2,\hat{V}_{\hat{\xi}}^T) \tag{4.14}$$

is satisfied.

Similarly, as in Chapter 3, we can obtain the corresponding analytical expression for $\dot{y}$ from (4.5) with the measurement noise set to zero, as

$$\dot{y}=\mathcal{L}_{\tilde{f}_1+\tilde{g}_{11}w}\tilde{h}_{21}+\mathcal{L}_{\tilde{f}_2+\tilde{g}_{21}w}\tilde{h}_{22},$$

which under certainty-equivalence and in the presence of $\hat{w}^\star$, results in

$$\begin{aligned}\dot{y} &= \mathcal{L}_{\tilde{f}_1(\hat{\xi}_1)++\tilde{g}_{11}(\hat{\xi})\hat{w}^\star}\tilde{h}_{21}(\hat{\xi}_1)+\mathcal{L}_{\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2)+\frac{1}{\varepsilon}\tilde{g}_{21}(\hat{\xi})\hat{w}^\star}\tilde{h}_{22}(\hat{\xi}_2),\\ &= \nabla_{\hat{\xi}_1}h_{21}(\hat{\xi}_1)[\tilde{f}_1(\hat{\xi}_1)+\frac{1}{\gamma^2}\tilde{g}_{11}(\hat{\xi})\tilde{g}_{11}^T(\hat{\xi})\hat{V}_{\hat{\xi}_1}^T(\hat{\xi},y)+\frac{1}{\gamma^2\varepsilon}\tilde{g}_{11}(\hat{\xi})\tilde{g}_{21}^T(\hat{\xi})\hat{V}_{\hat{\xi}_2}^T(\hat{\xi},y)]+\\ &\quad \nabla_{\hat{\xi}_2}h_{22}(\hat{\xi}_2)[\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2))+\frac{1}{\gamma^2\varepsilon}\tilde{g}_{21}(\hat{\xi})\tilde{g}_{11}^T(\hat{\xi})\hat{V}_{\hat{\xi}_1}^T(\hat{\xi},y)+\frac{1}{\gamma^2\varepsilon^2}\tilde{g}_{21}(\hat{\xi})\tilde{g}_{21}^T(\hat{\xi})\hat{V}_{\hat{\xi}_2}^T(\hat{\xi},y)] \quad (4.15)\end{aligned}$$

Finally, setting

$$\hat{H}(\hat{\xi},y,\hat{w}^\star,\hat{L}_1^\star,\hat{L}_2^\star,\hat{V}_{\hat{\xi}}^T,\hat{V}_y^T)=0$$

and using the above expression (4.15) for $\dot{y}$ results in the following Hamilton-Jacobi-Isaacs equation (HJIE):

$$\begin{aligned}
&\hat{V}_{\hat{\xi}_1}(\hat{\xi},y)\tilde{f}_1(\hat{\xi}_1)+\tfrac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi},y)\tilde{f}_2(\hat{\xi}_2)+\hat{V}_y(\hat{\xi},y)[\nabla_{\hat{\xi}_1}h_{21}(\hat{\xi}_1)\tilde{f}_1(\hat{\xi}_1)+\tfrac{1}{\varepsilon}\nabla_{\hat{\xi}_2}h_{22}(\hat{\xi}_2)\tilde{f}_2(\hat{\xi}_2)]+\\
&\tfrac{1}{2\gamma^2}[\hat{V}_{\hat{\xi}_1}(\hat{\xi},y)+2\hat{V}_y(\hat{\xi},y)\nabla_{\hat{\xi}_1}h_{21}(\hat{\xi}_1)\quad \hat{V}_{\hat{\xi}_2}(\hat{\xi},y)+2\hat{V}_y(\hat{\xi},y)\nabla_{\hat{\xi}_2}h_{22}(\hat{\xi}_2)]\times\\
&\begin{bmatrix}\tilde{g}_{11}(\hat{\xi})\tilde{g}_{11}^T(\hat{\xi}) & \tfrac{1}{\varepsilon}\tilde{g}_{11}(\hat{\xi})\tilde{g}_{21}^T(\hat{\xi})\\ \tfrac{1}{\varepsilon}\tilde{g}_{21}(\hat{\xi})\tilde{g}_{11}^T(\hat{\xi}) & \tfrac{1}{\varepsilon^2}\tilde{g}_{21}(\hat{\xi})\tilde{g}_{21}^T(\hat{\xi})\end{bmatrix}\begin{bmatrix}\hat{V}_{\hat{\xi}_1}^T(\hat{\xi},y)\\ \hat{V}_{\hat{\xi}_2}^T(\hat{\xi},y)\end{bmatrix}-\\
&\tfrac{1}{2}\hat{V}_{\hat{\xi}_1}(\hat{\xi},y)\hat{L}_1(\hat{\xi},y)\hat{L}_1^T(\hat{\xi},y)\hat{V}_{\hat{\xi}_1}^T(\hat{\xi},y)-\tfrac{1}{2\varepsilon^2}\hat{V}_{\hat{\xi}_2}(\hat{\xi},y)\hat{L}_2(\hat{\xi},y)\hat{L}_2^T(\hat{\xi},y)\hat{V}_{\hat{\xi}_2}^T(\hat{\xi},y)-\\
&\tfrac{1}{2}(y-\tilde{h}_{21}(\hat{\xi}_1)-\tilde{h}_{22}(\hat{\xi}_2))^T(y-\tilde{h}_{21}(\hat{\xi}_1)-\tilde{h}_{22}(\hat{\xi}_2))=0,\ \ \hat{V}(0,0)=0,
\end{aligned}\tag{4.16}$$

or equivalently the HJIE

$$\begin{aligned}
&\hat{V}_{\hat{\xi}_1}(\hat{\xi},y)\tilde{f}_1(\hat{\xi}_1)+\tfrac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi},y)\tilde{f}_2(\hat{\xi}_2)+\hat{V}_y(\hat{\xi},y)[\nabla_{\hat{\xi}_1}h_{21}(\hat{\xi}_1)\tilde{f}_1(\hat{\xi}_1)+\tfrac{1}{\varepsilon}\nabla_{\hat{\xi}_2}h_{22}(\hat{\xi}_2)\tilde{f}_2(\hat{\xi}_2)]+\\
&\tfrac{1}{2\gamma^2}[\hat{V}_{\hat{\xi}_1}(\hat{\xi},y)+2\hat{V}_y(\hat{\xi},y)\nabla_{\hat{\xi}_1}h_{21}(\hat{\xi}_1)\quad \hat{V}_{\hat{\xi}_2}(\hat{\xi},y)+2\hat{V}_y(\hat{\xi},y)\nabla_{\hat{\xi}_2}h_{22}(\hat{\xi}_2)]\times\\
&\begin{bmatrix}\tilde{g}_{11}(\hat{\xi})\tilde{g}_{11}^T(\hat{\xi}) & \tfrac{1}{\varepsilon}\tilde{g}_{11}(\hat{\xi})\tilde{g}_{21}^T(\hat{\xi})\\ \tfrac{1}{\varepsilon}\tilde{g}_{21}(\hat{\xi})\tilde{g}_{11}^T(\hat{\xi}) & \tfrac{1}{\varepsilon^2}\tilde{g}_{21}(\hat{\xi})\tilde{g}_{21}^T(\hat{\xi})\end{bmatrix}\begin{bmatrix}\hat{V}_{\hat{\xi}_1}^T(\hat{\xi},y)\\ \hat{V}_{\hat{\xi}_2}^T(\hat{\xi},y)\end{bmatrix}-\\
&\tfrac{3}{2}(y-\tilde{h}_{21}(\hat{\xi}_1)-\tilde{h}_{22}(\hat{\xi}_2))^T(y-\tilde{h}_{21}(\hat{\xi}_1)-\tilde{h}_{22}(\hat{\xi}_2))=0,\ \ \hat{V}(0,0)=0.
\end{aligned}\tag{4.17}$$

Moreover, from (4.10), we have

$$\begin{aligned}
\hat{H}(\hat{\xi},y,w,\hat{L}_1^\star,\hat{L}_2^\star,\hat{V}_{\hat{\xi}}^T,\hat{V}_y^T) &= \hat{V}_{\hat{\xi}_1}(\hat{\xi},y)\tilde{f}_1(\hat{\xi}_1)+\frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi},y)\tilde{f}_2(\hat{\xi}_2)+\hat{V}_y(\hat{\xi},y)\dot{y}-\\
&\quad \frac{\gamma^2}{2}\|w-w^\star\|^2+\frac{\gamma^2}{2}\|w^\star\|^2-\frac{3}{2}\|z\|^2\\
&= \hat{H}(\hat{\xi},y,w^\star,\hat{L}_1^\star,\hat{L}_2^\star,\hat{V}_{\hat{\xi}}^T,\hat{V}_y^T)-\frac{\gamma^2}{2}\|w-w^\star\|^2.
\end{aligned}$$

Thus,

$$\hat{H}(\hat{\xi},y,w,\hat{L}_1^\star,\hat{L}_2^\star,\hat{V}_{\hat{\xi}}^T,\hat{V}_y^T)\le \hat{H}(\hat{\xi},y,w^\star,\hat{L}_1^\star,\hat{L}_2^\star,\hat{V}_{\hat{\xi}}^T,\hat{V}_y^T).\tag{4.18}$$

Combining now (4.14) and (4.18), we have that the saddle-point conditions (4.9) are satisfied and the pair $([\hat{L}_1^\star,\hat{L}_2^\star],w^\star)$ constitutes a saddle-point solution to the game (4.8). Consequently, we have the following result.

Proposition 4.1.1. *Consider the nonlinear system (4.2) and the $\mathcal{H}_\infty$ local filtering problem for this system. Suppose the plant $\mathbf{P}^{\mathbf{a}}_{\mathbf{sp}}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and zero-input observable for all $\varepsilon \in [0, \varepsilon^\star)$. Further, suppose there exist a local diffeomorphism φ that transforms the system to the partially decoupled form (4.5), a C^1 positive-semidefinite function $\hat{V} : \hat{N} \times \hat{\Upsilon} \to \Re_+$ locally defined in a neighborhood $\hat{N} \times \hat{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\hat{\xi}, y) = (0, 0)$, and matrix functions $\hat{L}_i : \hat{N} \times \hat{\Upsilon} \to \Re^{n_i \times m}$, $i = 1, 2$, satisfying the HJIE (4.16) together with the side-conditions (4.12), (4.13) for some $\gamma > 0$ and $\varepsilon < \varepsilon^\star$ (that guarantees asymptotic stability of the system). Then, the filter $\mathbf{F}^a_{1c}$ solves the local $\mathcal{H}_\infty$ filtering problem for the system.*

Proof: The optimality of the filter gains $\hat{L}_1^\star$, $\hat{L}_2^\star$ has already been shown above. It remains to prove asymptotic convergence of the estimation error vector. Accordingly, let $\hat{V} \geq 0$ be a C^1 solution of the HJIE (4.16) or equivalently (4.17). Then, differentiating this solution along a trajectory of (4.6) with $\hat{L}_1 = \hat{L}_1^\star$, $L_2 = \hat{L}_2^\star$, and any $w \in \mathcal{W}$ inplace of $\hat{w}^\star$, we get

$$
\begin{aligned}
\dot{\hat{V}} &= \hat{V}_{\hat{\xi}_1}(\hat{\xi}, y)[\tilde{f}_1(\hat{\xi}_1) + \tilde{g}_{11}(\hat{\xi})w + \hat{L}_1^\star(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))] + \\
&\quad \frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\hat{\xi}, y)[\tilde{f}_2(\hat{\xi}_2) + \tilde{g}_{21}(\hat{\xi})w + \hat{L}_2^\star(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}_1) - \tilde{h}_{22}(\hat{\xi}_2))] + \hat{V}_y(\hat{\xi}, y)\dot{y} \\
&= -\frac{\gamma^2}{2}\|w - \hat{w}^\star\|^2 + \frac{1}{2}\gamma^2\|w\|^2 - \frac{1}{2}\|z\|^2 \\
&\leq \frac{1}{2}\gamma^2\|w\|^2 - \frac{1}{2}\|z\|^2,
\end{aligned}
$$

where the last equality follows from using the HJIE (4.16). Integrating the above inequality from $t = t_0$ to $t = \infty$ and since the system is asymptotically stable, implies that the $\mathcal{L}_2$-gain condition (4.3) is satisfied.

Moreover, setting $w = 0$ in the above inequality implies that $\dot{\hat{V}}(\hat{\xi}(t), y(t)) \leq -\frac{1}{2}\|z\|^2$. Therefore, the filter dynamics is stable, and $\hat{V}(\hat{\xi}(t), y(t))$ is non-increasing along a trajectory of (4.6). Further, the condition that $\dot{\hat{V}}(\hat{\xi}(t), y(t)) \equiv 0$ $\forall t \geq t_s$ implies that $z \equiv 0$, which further implies that $y = \tilde{h}_{21}(\hat{\xi}_1) + \tilde{h}_{22}(\hat{\xi}_2)$ $\forall t \geq t_s$. By the zero-input observability of the system, this implies that $\hat{\xi} = \xi$. Finally, since φ is invertible and $\varphi(0) = 0$, $\hat{\xi} = \xi$ implies $\hat{x} = \varphi^{-1}(\hat{\xi}) = \varphi^{-1}(\xi) = x$. □

Remark 4.1.1. *Note that, we have not included the term* $k_{21}(\hat{\xi})\hat{w}^\star$ *as part of the innovation variable (or estimation error),* $e = y - h_{21}(\hat{\xi}_1) - h_{22}(\hat{\xi}_2)$*, in the filter design (4.6) only to simplify the design. Moreover, the benefit of including it is very marginal.*

Remark 4.1.2. *An estimation of the upper-bound* ε^* *of the singular perturbation parameter* ε *that guarantees the asymptotic stability of the filter (4.6) and the satisfaction of the* $\mathcal{L}_2$*-gain condition (4.3) can be made from a local linearization about* $\hat{\xi} = 0$ *and a linear analysis of the filter (4.6). This will be discussed after Corollary 4.1.1.*

To relate the above result to the linear theory (Gajic, 1994), (Haddad, 1976), we consider the following linear singularly-perturbed system (LSPS):

$$\mathbf{P}^l_{sp} : \begin{cases} \dot{x}_1 &= A_1x_1 + A_{12}x_2 + B_{11}w; \quad x_1(t_0) = x_{10} \\ \varepsilon\dot{x}_2 &= A_{21}x_1 + A_2x_2 + B_{21}w; \quad x_2(t_0) = x_{20} \\ y &= C_{21}x_1 + C_{22}x_2 + w \end{cases} \tag{4.19}$$

where $A_1 \in \Re^{n_1\times n_1}$, $A_{12} \in \Re^{n_1\times n_2}$, $A_{21} \in \Re^{n_2\times n_1}$, $A_2 \in \Re^{n_2\times n_2}$, $B_{11} \in \Re^{n_1\times s}$, and $B_{21} \in \Re^{n_2\times s}$, while the other matrices have compatible dimensions. Then, an explicit form of the required transformation φ above is given by the Chang transformation (Chang, 1972):

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} I_{n_1} - \varepsilon\mathsf{HL} & -\varepsilon\mathsf{H} \\ \mathsf{L} & I_{n_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4.20}$$

where the matrices L and H satisfy the equations

$$\begin{aligned} 0 &= A_2\mathsf{L} - A_{21} - \varepsilon\mathsf{L}(A_1 - A_{12}\mathsf{L}) \\ 0 &= -\mathsf{H}(A_2 + \varepsilon\mathsf{L}A_{12}) + A_{12} + \varepsilon(A_1 - A_{12}\mathsf{L})\mathsf{H}. \end{aligned}$$

The system is then represented in the new coordinates by

$$\tilde{\mathbf{P}}^l_{sp} : \begin{cases} \dot{\xi}_1 &= \tilde{A}_1\xi_1 + \tilde{B}_{11}w; \quad \xi_1(t_0) = \xi_{10} \\ \varepsilon\dot{\xi}_2 &= \tilde{A}_2\xi_2 + \tilde{B}_{21}w; \quad \xi_2(t_0) = \xi_{20} \\ y &= \tilde{C}_{21}\xi_1 + \tilde{C}_{22}\xi_2 + w \end{cases} \tag{4.21}$$

where

$$\begin{aligned}
\tilde{A}_1 &= A_1 - A_{12}\mathsf{L} = A_1 - A_{12}A_2^{-1}A_{21} + O(\varepsilon)\\
\tilde{B}_{11} &= B_{11} - \varepsilon\mathsf{H}\mathsf{L}B_{11} - \mathsf{H}B_{21} = B_{11} - A_{12}A_2^{-1}B_{21} + O(\varepsilon)\\
\tilde{A}_2 &= A_2 + \varepsilon\mathsf{L}A_{12} = A_2 + O(\varepsilon)\\
\tilde{B}_{21} &= B_{21} + \varepsilon\mathsf{L}B_{11} = B_{21} + O(\varepsilon)\\
\tilde{C}_{21} &= C_{21} - C_{22}\mathsf{L} = C_{21} - C_{22}A_2^{-1}A_{21} + O(\varepsilon)\\
\tilde{C}_{22} &= C_{22} + \varepsilon(C_{21} - C_{22})\mathsf{H} = C_{22} + O(\varepsilon).
\end{aligned}$$

Adapting the filter (4.6) to the system (4.21) then yields the following filter

$$\mathbf{F}^l_{1c}: \begin{cases} \dot{\hat{\xi}}_1 &= (\tilde{A}_1 + \frac{1}{\gamma^2}\tilde{B}_{11}\tilde{B}_{11}^T\hat{P}_1)\hat{\xi}_1 + \frac{1}{\gamma^2\varepsilon}\tilde{B}_{11}\tilde{B}_{21}^T\hat{P}_2\hat{\xi}_2 + \hat{L}_1(y - \tilde{C}_{21}\hat{\xi}_1 - \tilde{C}_{22}\hat{\xi}_2),\\ &\quad \hat{\xi}_1(t_0) = 0\\ \varepsilon\dot{\hat{\xi}}_2 &= (\tilde{A}_2 + \frac{1}{\gamma^2\varepsilon}\tilde{B}_{21}\tilde{B}_{21}^T\hat{P}_2)\hat{\xi}_2 + \frac{1}{\gamma^2}\tilde{B}_{21}\tilde{B}_{11}^T\hat{P}_1\hat{\xi}_1 + \hat{L}_2(y - \tilde{C}_{21}\hat{\xi}_1 - \tilde{C}_{22}\hat{\xi}_2),\\ &\quad \hat{\xi}_2(t_0) = 0, \end{cases} \tag{4.22}$$

where $\hat{P}_1$, $\hat{P}_2$, $\hat{L}_1$, $\hat{L}_2$ satisfy the following matrix inequalities:

$$\begin{bmatrix} \tilde{A}_1^T\hat{P}_1 + \hat{P}_1\tilde{A}_1 + \frac{1}{\gamma^2}\hat{P}_1\tilde{B}_{11}\tilde{B}_{11}^T\hat{P}_1 - 3\tilde{C}_{21}^T\tilde{C}_{21} & \frac{1}{\gamma^2\varepsilon}\hat{P}_1\tilde{B}_{11}\tilde{B}_{21}^T\hat{P}_2 + 3\tilde{C}_{21}^T\tilde{C}_{22} & 3\tilde{C}_{21}^T & 0\\ \frac{1}{\gamma^2\varepsilon}\hat{P}_2\tilde{B}_{21}\tilde{B}_{11}^T\hat{P}_1 + 3\tilde{C}_{22}^T\tilde{C}_{21} & \tilde{A}_2^T\hat{P}_2 + \hat{P}_2\tilde{A}_2 + \frac{1}{\gamma^2\varepsilon}\hat{P}_2\tilde{B}_{21}\tilde{B}_{21}^T\hat{P}_2 - 3\tilde{C}_{22}^T\tilde{C}_{22} & 3\tilde{C}_{22}^T & 0\\ 3\tilde{C}_{21} & 3\tilde{C}_{22} & -3I & \frac{1}{2}Q\\ 0 & 0 & \frac{1}{2}Q & 0 \end{bmatrix} \le 0, \tag{4.23}$$

$$\begin{bmatrix} 0 & 0 & \frac{1}{2}(\hat{P}_1\hat{L}_1 - \tilde{C}_{21}^T)\\ 0 & 0 & -\frac{1}{2}\tilde{C}_{22}^T\\ \frac{1}{2}(\hat{P}_1\hat{L}_1 - \tilde{C}_{21}^T)^T & -\frac{1}{2}\tilde{C}_{22}^T & (1-\mu_1)I \end{bmatrix} \le 0, \tag{4.24}$$

$$\begin{bmatrix} 0 & 0 & -\frac{1}{2}\tilde{C}_{21}^T \\ 0 & 0 & \frac{1}{2\varepsilon}(\hat{P}_2\hat{L}_2 - \tilde{C}_{22}^T) \\ -\frac{1}{2}\tilde{C}_{21} & \frac{1}{2\varepsilon}(\hat{P}_2\hat{L}_2 - \tilde{C}_{22}^T)^T & (1-\mu_2)I \end{bmatrix} \leq 0, \tag{4.25}$$

for some symmetric matrix $Q \in \Re^{m\times m} \geq 0$, and numbers $\mu_1, \mu_2 \geq 1$. Consequently, we have the following Corollary to Proposition 4.1.1.

Corollary 4.1.1. *Consider the linear system (4.19) and the $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\mathbf{P}_{sp}^l$ is locally asymptotically stable about the equilibrium-point $x = 0$ and observable for all $\varepsilon \in [0, \varepsilon^\star)$. Suppose further, it is transformable to the form (4.21), and there exist positive-semidefinite matrices $\hat{P}_1 \in \Re^{n_1\times n_1}$, $\hat{P}_2 \in \Re^{n_2\times n_2}$, $Q \in \Re^{m\times m}$, and matrices $\hat{L}_1, \hat{L}_2 \in \Re^{n\times m}$, satisfying the matrix-inequalities (MIs) (4.23)-(4.25) for some $\gamma > 0$ and $\varepsilon < \varepsilon^\star$. Then the filter $\mathbf{F}_{1c}^l$ solves the $\mathcal{H}_\infty$ filtering problem for the system.*

Proof: Take

$$\hat{V}(\hat{\xi}, y) = \frac{1}{2}(\hat{\xi}_1^T P_1 \hat{\xi}_1 + \hat{\xi}_2^T P_2 \hat{\xi}_2 + y^T Q y)$$

and apply the result of the Proposition. □

Furthermore, to estimate an upper-bound ε^* on the singular perturbation parameter ε that guarantees the asymptotic stability of the filter (4.6) and the satisfaction of the $\mathcal{L}_2$-gain condition (4.3), the result of the above Corollary 4.1.1 can be utitilized to formulate an optimization problem. By assuming that the model (4.19), is a local linearization about $x = 0$ of the nonlinear model (4.2) in the sense that,

$$\left.\begin{array}{lll} A_1 = \frac{\partial f_1}{\partial x_1}\Big|_{x=0}(x_1,x_2), & A_{12} = \frac{\partial f_1}{\partial x_2}\Big|_{x=0}(x_1,x_2), & B_{11} = g_{11}(0,0), \\ A_{21} = \frac{\partial f_2}{\partial x_1}\Big|_{x=0}(x_1,x_2), & A_2 = \frac{\partial f_2}{\partial x_2}\Big|_{x=0}(x_1,x_2), & B_{21} = g_{21}(0,0), \\ C_{21} = \frac{\partial h_{21}}{\partial x_1}\Big|_{x=0}(x_1,x_2), & C_{22} = \frac{\partial h_{22}}{\partial x_2}\Big|_{x=0}(x_1,x_2), & k_{21}(0,0) = I, \end{array}\right\} \tag{4.26}$$

we can then state the following corollary.

Corollary 4.1.2. *Consider the nonlinear system (4.2) and the $\mathcal{H}_\infty$ filtering problem for this system. Let (4.26) be a local linearization of the system, and suppose the system is locally asymptotically stable about the equilibrium-point $x = 0$ for all $\varepsilon \in [0, \varepsilon^\star)$ and zero-input* *observable. Suppose further, it is transformable to the form (4.21), and there exist positive-semidefinite matrices* $\hat{P}_1 \in \Re^{n_1 \times n_1}$*,* $\hat{P}_2 \in \Re^{n_2 \times n_2}$*,* $Q \in \Re^{m \times m}$*, matrices* $\hat{L}_1, \hat{L}_2 \in \Re^{n \times m}$*, and numbers* γ^**,* ε^* *that solve the optimization problem:*

$$\min \gamma - \varepsilon \quad s.t. \quad (4.23) - (4.25). \tag{4.27}$$

Then, the filter $\mathbf{F}^l_{1c}$ *solves the* $\mathcal{H}_\infty$ *filtering problem for the system locally. Moreover,* γ^* *is the minimum achievable disturbance attenuation level for the filter, and* ε^* *is an upper-bound of the parameter* ε *for asymptotic stability of the filter.*

Remark 4.1.3. *Notice, in the above Corollary 4.1.2, it is possible to have* $\varepsilon^* \geq \varepsilon^\star$*.*

Proposition 4.1.1 has not yet exploited the benefit of the coordinate transformation in designing the filter (4.6) for the system (4.5). Moreover, for the linear system (4.19), the resulting governing equations (4.23)-(4.25) are not linear in the unknown variables $\hat{P}_1$, $\hat{P}_2$. Thus, we shall now design separate reduced-order filters for the decomposed subsystems which should be more efficient than the previous one. For this purpose, we let $\varepsilon \downarrow 0$ in (4.5) and obtain the following reduced system model:

$$\widetilde{\mathbf{P}}^{\mathbf{a}}_{\mathbf{r}} : \begin{cases} \dot{\xi}_1 &= \tilde{f}_1(\xi_1) + \tilde{g}_{11}(\xi)w \\ 0 &= \tilde{f}_2(\xi_2) + \tilde{g}_{21}(\xi)w \\ y &= \tilde{h}_{21}(\xi_1) + \tilde{h}_{22}(\xi_2) + \tilde{k}_{21}(\xi)w. \end{cases} \tag{4.28}$$

Then we assume the following.

Assumption 4.1.1. *The system (4.2), (4.28) is in the "standard form", i.e., the equation*

$$0 = \tilde{f}_2(\xi_2) + \tilde{g}_{21}(\xi)w \tag{4.29}$$

has $l \geq 1$ *isolated roots, we can denote any one of these solutions by*

$$\bar{\xi}_2 = q(\xi_1, w). \tag{4.30}$$

for some smooth function $q : \mathcal{X} \times \mathcal{W} \to \mathcal{X}$*.*

Under Assumption 4.1.1, we obtain the reduced-order slow subsystem

$$\mathbf{P_r^a}:\begin{cases}\dot{\xi}_1 &=& \tilde{f}_1(\xi_1)+\tilde{g}_{11}(\xi_1,q(\xi_1,w))w+O(\varepsilon)\\ y &=& \tilde{h}_{21}(\xi_1)+\tilde{h}_{22}(q(\xi_1,w))+\tilde{k}_{21}(\xi_1,q(\xi_1,w))w+O(\varepsilon)\end{cases} \tag{4.31}$$

and a boundary-layer (or quasi-steady-state) subsystem as

$$\frac{d\bar{\xi}_2}{d\tau}=\tilde{f}_2(\bar{\xi}_2(\tau))+\tilde{g}_{21}(\xi_1,\bar{\xi}_2(\tau))w \tag{4.32}$$

where $\tau = t/\varepsilon$ is a stretched-time parameter. It can be shown that there exists an $\varepsilon^\star > 0$, such that this subsystem is asymptotically stable for all $\varepsilon \in (0, \varepsilon^\star)$ (see Theorem 8.2 in Ref. (Khalil, 1985)) if the original system (4.2) is asymptotically stable.

We can therefore proceed to redesign the filter in (4.6) for the composite system (4.31), (4.32) separately as

$$\widetilde{\mathbf{F}}^a_{2c}:\begin{cases}\dot{\breve{\xi}}_1 &=& \tilde{f}_1(\breve{\xi}_1)+\tilde{g}_{11}(\breve{\xi}_1,q(\breve{\xi}_1,\breve{w}_1^\star))\breve{w}_1^\star+\breve{L}_1(\breve{\xi}_1,y)(y-\tilde{h}_{21}(\breve{\xi}_1)-\\ && h_{22}(q(\breve{\xi}_1,\breve{w}_1^\star))),\quad \breve{\xi}_1(t_0)=0\\ \varepsilon\dot{\breve{\xi}}_2 &=& \tilde{f}_2(\breve{\xi}_2)+\tilde{g}_{21}(\breve{\xi})w_2^\star+\breve{L}_2(\breve{\xi}_2,y)(y-\tilde{h}_{21}(\breve{\xi}_1)-\tilde{h}_{22}(\breve{\xi}_2)),\;\; \breve{\xi}_2(t_0)=0.\\ \breve{z} &=& y-\tilde{h}_{21}(\breve{\xi}_1)-\tilde{h}_{22}(\breve{\xi}_2),\end{cases} \tag{4.33}$$

where we have decomposed w into two components w_1 and w_2 for convenience, and $\breve{w}_i^\star$ is predetermined with $\breve{\xi}_j$ constant (Chow, 1976), $i \neq j,\ i,j = 1,2$. Notice also that, ξ_2 cannot be estimated from (4.30) since this is a "quasi-steady-state" approximation.

The following theorem then summaries this design approach.

Theorem 4.1.1. *Consider the nonlinear system (4.2) and the $\mathcal{H}_\infty$ local filtering problem for this system. Suppose the plant $\mathbf{P^a_{sp}}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and zero-input observable for all $\varepsilon \in [0, \varepsilon^\star)$. Further, suppose there exist a local diffeomorphism φ that transforms the system to the partially decoupled form (4.5), and Assumption 4.1.1 holds. In addition, suppose for some $\gamma > 0$ and $\varepsilon \in [0, \varepsilon^\star)$, there exist C^1 positive-semidefinite functions $\breve{V}_i : \breve{N}_i \times \breve{\Upsilon}_i \to \Re_+,\ i = 1,2,$ locally defined in neighborhoods*

$\breve{N}_i \times \breve{\Upsilon}_i \subset \mathcal{X} \times \mathcal{Y}$ *of the origin* $(\breve{\xi}_i, y) = (0,0)$, $i = 1,2$ *respectively, and matrix functions* $\breve{L}_i : \breve{N}_i \times \breve{\Upsilon}_i \to \Re^{n_i \times m}$, $i = 1,2$ *satisfying the HJIEs:*

$$\begin{gathered}\breve{V}_{1\breve{\xi}_1}(\breve{\xi}_1, y)\tilde{f}_1(\breve{\xi}_1) + \tfrac{1}{2\gamma^2}\breve{V}_{1\breve{\xi}_1}(\breve{\xi}_1, y)\tilde{g}_{11}(\breve{\xi}_1, q(\breve{\xi}_1, \breve{w}_1^\star))\tilde{g}_{11}^T(\breve{\xi}_1, q(\breve{\xi}_1, \breve{w}_1^\star))\breve{V}_{1\breve{\xi}_1}^T(\breve{\xi}_1, y) + \\ \breve{V}_{1y}(\breve{\xi}_1, y)[\nabla_{\breve{\xi}_1}\tilde{h}_{21}(\breve{\xi}_1) + \nabla_{\breve{\xi}_1}\tilde{h}_{22}(q(\breve{\xi}_1, \breve{w}_1^\star))]\Big(\tilde{f}_1(\breve{\xi}_1) + \\ \tfrac{1}{\gamma^2}\tilde{g}_{11}(\breve{\xi}_1, q(\breve{\xi}_1, \breve{w}_1^\star))\tilde{g}_{11}^T(\breve{\xi}_1, q(\breve{\xi}_1, \breve{w}_1^\star))\breve{V}_{1\breve{\xi}_1}^T(\breve{\xi}_1, y)\Big) - \tfrac{1}{2}(y - \tilde{h}_{21}(\breve{\xi}_1) - \\ \tilde{h}_{22}(q(\breve{\xi}_1, \breve{w}_1^\star))^T(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(q(\breve{\xi}_1, \breve{w}_1^\star))) = 0, \quad \breve{V}_1(0,0) = 0\end{gathered} \tag{4.34}$$

$$\begin{gathered}\tfrac{1}{\varepsilon}\breve{V}_{2\breve{\xi}_2}(\breve{\xi}, y)\tilde{f}_2(\breve{\xi}_2) + \tfrac{1}{2\gamma^2\varepsilon^2}\breve{V}_{2\breve{\xi}_2}(\breve{\xi}, y)\tilde{g}_{21}(\breve{\xi})\tilde{g}_{21}^T(\breve{\xi})\breve{V}_{2\breve{\xi}_2}^T(\breve{\xi}, y) + \\ \breve{V}_{2y}(\breve{\xi}_2, y)\Big([\tfrac{1}{\varepsilon}\nabla_{\breve{\xi}_2}\tilde{h}_{22}(\breve{\xi}_2)\tilde{f}_2(\breve{\xi}_2) + \tfrac{1}{\gamma^2\varepsilon^2}\tilde{g}_{21}(\breve{\xi})\tilde{g}_{21}^T(\breve{\xi})\breve{V}_{2\breve{\xi}_2}^T(\breve{\xi}, y)] + \\ [\nabla_{\breve{\xi}_1}\tilde{h}_{21}(\breve{\xi}_1)\tilde{f}_1(\breve{\xi}_1) + \tfrac{1}{\gamma^2}\tilde{g}_{11}(\breve{\xi}_1, q(\breve{\xi}_1, \breve{w}_1^\star))\tilde{g}_{11}^T(\breve{\xi}_1, q(\breve{\xi}_1, \breve{w}_1^\star))\breve{V}_{1\breve{\xi}_1}^T(\breve{\xi}_1, y)]\Big) - \\ \tfrac{1}{2}(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(\breve{\xi}_2))^T(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(\breve{\xi}_2)) = 0, \quad \breve{V}_2(0,0) = 0\end{gathered} \tag{4.35}$$

and where

$$\breve{w}_1^\star = \frac{1}{\gamma^2}\tilde{g}_{11}^T(\breve{\xi}_1, \bar{\xi}_2)\breve{V}_{1\breve{\xi}_1}^T(\breve{\xi}_1, y),$$

together with the side-conditions

$$\breve{V}_{1\breve{\xi}_1}(\breve{\xi}_1, y)\breve{L}_1(\breve{\xi}_1, y) = -(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(q(\breve{\xi}_1, \breve{w}_1^\star)))^T \tag{4.36}$$

$$\breve{V}_{2\breve{\xi}_2}(\breve{\xi}, y)\breve{L}_2(\breve{\xi}, y) = -\varepsilon(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(\breve{\xi}_2))^T. \tag{4.37}$$

Then, the filter $\widetilde{\mathbf{F}}^{\mathbf{a}}_{\mathbf{2c}}$ *solves the local* $\mathcal{H}_\infty$ *filtering problem for the system.*

Proof: (Sketch). We define separately two Hamiltonian functions $\breve{H}_i : T^\star\mathcal{X} \times \mathcal{W} \times \Re^{n_i \times m} \to \Re$, $i = 1,2$ with respect to the cost-functional (4.8) for each of the two separate components of the filter (4.33) as

$$\begin{aligned}\breve{H}_1(\breve{\xi}_1, y, w_1, \breve{L}_1, \breve{L}_2, \breve{V}_{\breve{\xi}_1}^T, \breve{V}_y^T) &= \breve{V}_{1\breve{\xi}_1}(\breve{\xi}_1, y)[\tilde{f}_1(\breve{\xi}_1) + \tilde{g}_{11}(\breve{\xi}_1, \bar{\xi}_2)w_1 + \breve{L}_1(\breve{\xi}_1, y)(y - \\ &\quad \tilde{h}_{21}(\breve{\xi}_1) - h_{22}(\bar{\xi}_2))] + \frac{1}{2}(\|z\|^2 - \gamma^2\|w_1\|^2)\end{aligned} \tag{4.38}$$

$$\breve{H}_2(\breve{\xi}, y, w_2, \breve{L}_1, \breve{L}_2, \breve{V}_{\breve{\xi}}^T, \breve{V}_y^T) \;=\; \frac{1}{\varepsilon}\breve{V}_{2\breve{\xi}_2}(\breve{\xi}, y)[\tilde{f}_2(\breve{\xi}_2) + \tilde{g}_{21}(\breve{\xi})w_2 + \breve{L}_2(\breve{\xi}, y)(y - \tilde{h}_{21}(\breve{\xi}_1) - \tilde{h}_{22}(\breve{\xi}_2))] + \frac{1}{2}(\|z\|^2 - \gamma^2\|w_2\|^2) \qquad (4.39)$$

for some smooth functions $\breve{V}_i : \mathcal{X} \times \mathcal{Y} \to \Re$, $i = 1, 2$. Then, we can determine $\breve{w}_1^\star$, $\breve{w}_2^\star$ by applying the necessary conditions for the worst-case noise as

$$\begin{aligned} \breve{w}_1^\star &= \frac{1}{\gamma^2}\tilde{g}_{11}^T(\breve{\xi}_1, \bar{\xi}_2)\breve{V}_{1\breve{\xi}_1}^T(\breve{\xi}_1, y) \\ \breve{w}_2^\star &= \frac{1}{\varepsilon\gamma^2}\tilde{g}_{12}^T(\breve{\xi})\breve{V}_{2\breve{\xi}_2}^T(\breve{\xi}, y) \end{aligned}$$

where $\breve{w}_1^\star$ is determined with $\bar{\xi}_2$ fixed. The rest of the proof follows along the same lines as Proposition 4.1.1. □

The limiting behavior of the filter (4.33) as $\varepsilon \downarrow 0$ corresponds to the reduced-order filter

$$\widetilde{\mathbf{F}}_{2r}^a : \begin{cases} \dot{\breve{\xi}}_1 = \tilde{f}_1(\breve{\xi}_1) + \tilde{g}_{11}(\breve{\xi}_1, q(\breve{\xi}_1, \breve{w}_1^\star))\breve{w}_1^\star + \breve{L}_1(\breve{\xi}_1, y)[y - \tilde{h}_{21}(\breve{\xi}_1) - \\ \qquad h_{22}(q(\breve{\xi}_1, \breve{w}_1^\star))], \quad \breve{\xi}_1(t_0) = 0, \end{cases} \qquad (4.40)$$

which is governed by the HJIE (4.34).

Similarly, specializing the result of Theorem 4.1.1 to the linear system (4.19). Assuming A_2 is nonsingular (Assumption 4.1.1), we have

$$\bar{\xi}_2 = -A_2^{-1}B_{21}w,$$

and hence we obtain the composite filter

$$\mathbf{F}_{2c}^l : \begin{cases} \dot{\breve{\xi}}_1 = \tilde{A}_1\breve{\xi}_1 + \frac{1}{\gamma^2}\tilde{B}_{11}\tilde{B}_{11}^T\breve{P}_1\breve{\xi}_1 + \breve{L}_1(y - \tilde{C}_{21}\breve{\xi}_1 + \frac{1}{\gamma^2}\tilde{C}_{22}\tilde{A}_2^{-1}\tilde{B}_{21}\tilde{B}_{11}^T\breve{P}_1\breve{\xi}_1), \\ \qquad \breve{\xi}_1(t_0) = 0 \\ \varepsilon\dot{\breve{\xi}}_2 = \tilde{A}_2\breve{\xi}_2 + \frac{1}{\gamma^2\varepsilon}\tilde{B}_{21}\tilde{B}_{21}^T\breve{P}_2\xi_2 + \breve{L}_2(y - \tilde{C}_{21}\breve{\xi}_1 - \tilde{C}_{22}\breve{\xi}_2), \\ \qquad \breve{\xi}_2(t_0) = 0. \end{cases} \qquad (4.41)$$

The following corollary summarizes this development.

Corollary 4.1.3. *Consider the linear system (4.19) and the $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\mathbf{P}^l_{sp}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and observable for all $\varepsilon \in [0, \varepsilon^\star)$. Suppose further, it is transformable to the form (4.21) and Assumption 4.1.1 holds or A_2 is nonsingular. In addition, suppose for some $\gamma > 0$ and $\varepsilon \in [0, \varepsilon^\star)$, there exist positive-semidefinite matrices $\breve{P}_1 \in \Re^{n_1 \times n_1}$, $\breve{P}_2 \in \Re^{n_2 \times n_2}$, $\breve{Q}_1$, $\breve{Q}_2 \in \Re^{m \times m}$ and matrices $\breve{L}_1 \in \Re^{n_1 \times m}$, $\breve{L}_2 \in \Re^{n_2 \times m}$, satisfying the linear-matrix-inequalities (LMIs)*

$$\begin{bmatrix} \begin{pmatrix} \tilde{A}_1^T \breve{P}_1 + \breve{P}_1 \tilde{A}_1 - \tilde{C}_{21}^T \tilde{C}_{21} + \frac{1}{\gamma^2} \tilde{C}_{21}^T \tilde{C}_{22} \tilde{A}_2^{-1} \tilde{B}_{21} \tilde{B}_{11}^T \breve{P}_1 + \\ \frac{1}{\gamma^2} \breve{P}_1 \tilde{B}_{11} \tilde{B}_{21}^T \tilde{A}_2^{-T} \tilde{C}_{22}^T \tilde{C}_{21} \end{pmatrix} & \breve{P}_1 \tilde{B}_{11} & \frac{1}{\gamma^2} \tilde{C}_{22} \tilde{A}_2^{-1} \tilde{B}_{21} \tilde{B}_{11}^T \breve{P}_1 & \tilde{C}_{21}^T - \frac{1}{\gamma^2} \breve{P}_1 \tilde{B}_{11} \tilde{B}_{21}^T \tilde{A}_2^{-T} \tilde{C}_{22}^T & 0 \\ \tilde{B}_{11}^T \breve{P}_1 & -\gamma^{-2} I & 0 & 0 & 0 \\ \breve{P}_1 \tilde{B}_{11} \tilde{B}_{21}^T \tilde{A}_2^{-T} \tilde{C}_{22}^T & 0 & -\gamma^{-2} I & 0 & 0 \\ \tilde{C}_{21} - \frac{1}{\gamma^2} \tilde{C}_{22} \tilde{A}_2^{-1} \tilde{B}_{21} \tilde{B}_{11}^T \breve{P}_1 & 0 & 0 & -I & \breve{Q}_1 \\ 0 & 0 & 0 & \breve{Q}_1 & 0 \end{bmatrix} \le 0 \qquad (4.42)$$

$$\begin{bmatrix} -\tilde{C}_{21}^T \tilde{C}_{21} & -\tilde{C}_{21}^T \tilde{C}_{22} & 0 & \tilde{C}_{21}^T & 0 \\ -\tilde{C}_{22}^T \tilde{C}_{21} & \frac{1}{\varepsilon} (\tilde{A}_2^T \breve{P}_2 + \breve{P}_2 \tilde{A}_2) - \tilde{C}_{22}^T \tilde{C}_{21} & \breve{P}_2 \tilde{B}_{21} & \tilde{C}_{22}^T & 0 \\ 0 & \tilde{B}_{21}^T \breve{P}_2 & -\varepsilon^{-2} \gamma^{-2} I & 0 & 0 \\ \tilde{C}_{21} & \tilde{C}_{22} & 0 & -I & \breve{Q}_2 \\ 0 & 0 & 0 & \breve{Q}_2 & 0 \end{bmatrix} \le 0 \qquad (4.43)$$

$$\begin{bmatrix} 0 & \begin{pmatrix} \frac{1}{2}(\breve{P}_1\breve{L}_1 - \tilde{C}_{21}^T + \\ \frac{1}{\gamma^2}\breve{P}_1\tilde{B}_{11}\tilde{B}_{21}\tilde{A}_2^{-T}\tilde{C}_{22}^T) \end{pmatrix} \\ \begin{pmatrix} \frac{1}{2}(\breve{P}_1\breve{L}_1 - \tilde{C}_{21}^T + \\ \frac{1}{\gamma^2}\breve{P}_1\tilde{B}_{11}\tilde{B}_{21}\tilde{A}_2^{-T}\tilde{C}_{22}^T)^T \end{pmatrix} & (1-\delta_1)I \end{bmatrix} \leq 0 \tag{4.44}$$

$$\begin{bmatrix} 0 & 0 & -\frac{1}{2}\tilde{C}_{21}^T \\ 0 & 0 & \frac{1}{2}(\frac{1}{\varepsilon}\breve{P}_2\breve{L}_2 - \tilde{C}_{22}^T) \\ -\frac{1}{2}\tilde{C}_{21} & \frac{1}{2}(\frac{1}{\varepsilon}\breve{P}_2\breve{L}_2 - \tilde{C}_{22}^T)^T & (1-\delta_2)I \end{bmatrix} \leq 0 \tag{4.45}$$

for some numbers $\delta_1, \delta_2 \geq 1$. Then the filter $\mathbf{F}_{2c}^l$ solves the $\mathcal{H}_\infty$ filtering problem for the system.

Proof: Take

$$\begin{aligned} \breve{V}_1(\breve{\xi}_1, y) &= \frac{1}{2}(\breve{\xi}_1^T\breve{P}_1\breve{\xi}_1 + y^T\breve{Q}_1 y) \\ \breve{V}_2(\breve{\xi}_2, y) &= \frac{1}{2}(\breve{\xi}_2^T\breve{P}_2\breve{\xi}_2 + y^T\breve{Q}_2 y) \end{aligned}$$

and apply the result of the Theorem. Moreover, the nonsingularity of A_2 guarantees that a reduced-order subsystem exists. □

Remark 4.1.4. *A similar result to Corollary 4.1.2 can be obtained for the filter (4.33) in terms of the a local linearization about $x = 0$ represented by the filter (4.41) and based on the result of Corollary 4.1.2, to obtain an upper ε^* and lower bound γ^* for ε and γ respectively.*

4.1.3 Aggregate $\mathcal{H}_\infty$ Filters

In the absence of the coordinate transformation, φ, discussed in the previous Subsection, a filter has to be designed to solve the problem for the aggregate system (4.2). We discuss this class of filters in this subsection. Accordingly, consider the following class of filters:

$$\mathbf{F}^a_{3ag}:\begin{cases}\dot{\grave{x}}_1 &= f_1(\grave{x}_1,\grave{x}_2)+g_{11}(\grave{x}_1,\grave{x}_2)\grave{w}^\star+\grave{L}_1(\grave{x},y)(y-h_{21}(\grave{x}_1)+h_{22}(\grave{x}_2));\\ & \grave{x}_1(t_0)=0\\ \varepsilon\dot{\grave{x}}_2 &= f_2(\grave{x}_1,\grave{x}_2)+g_{12}(\grave{x}_1,\grave{x}_2)\grave{w}^\star+\grave{L}_2(\grave{x},y)(y-h_{21}(\grave{x}_1)+h_{22}(\grave{x}_2));\\ & \grave{x}_2(t_0)=0\\ \grave{z} &= y-h_{21}(\grave{x}_1)+h_{22}(\grave{x}_2),\end{cases}\tag{4.46}$$

where $\grave{L}_1\in\Re^{n_1\times m}$, $\grave{L}_2\in\Re^{n_2\times m}$ are the filter gains, and $\grave{z}$ is the new penalty variable. We can repeat the same kind of derivation above to arrive at the following.

Theorem 4.1.2. *Consider the nonlinear system (4.2) and the $\mathcal{H}_\infty$ local filtering problem for this system. Suppose the plant $\mathbf{P^a_{sp}}$ is locally asymptotically stable about the equilibrium-point $x=0$ and zero-input observable for all $\varepsilon\in[0,\varepsilon^\star)$. Further, suppose for some $\gamma>0$ and $\varepsilon\in[0,\varepsilon^\star)$, there exist a C^1 positive-semidefinite function $\grave{V}:\grave{N}\times\grave{\Upsilon}\to\Re_+$, locally defined in a neighborhood $\grave{N}\times\grave{\Upsilon}\subset\mathcal{X}\times\mathcal{Y}$ of the origin $(\grave{x}_1,\grave{x}_2,y)=(0,0,0)$, and matrix functions $\grave{L}_i:\grave{N}\times\grave{\Upsilon}\to\Re^{n_i\times m}$, $i=1,2$, satisfying the HJIE:*

$$\begin{aligned}&\grave{V}_{\grave{x}_1}(\grave{x},y)f_1(\grave{x}_1,\grave{x}_2)+\tfrac{1}{\varepsilon}\grave{V}_{\grave{x}_2}(\grave{x},y)f_2(\grave{x}_1,\grave{x}_2)+\grave{V}_y(\grave{x},y)[\nabla_{\grave{x}_1}h_{21}(\grave{x}_1)f_1(\grave{x})+\tfrac{1}{\varepsilon}\nabla_{\grave{x}_2}h_{22}(\grave{x}_2)f_2(\grave{x})]\\ &+\tfrac{1}{2\gamma^2}[\grave{V}_{\grave{x}_1}(\grave{x},y)+2\grave{V}_y(\grave{x},y)(\nabla_{\grave{x}_1}h_{21}(\grave{x}_1)\quad \grave{V}_{\grave{x}_2}(\grave{x},y)+2\nabla_{\grave{x}_2}h_{22}(\grave{x}_2)]\times\\ &\begin{bmatrix} g_{11}(\grave{x})g_{11}^T(\grave{x}) & \frac{1}{\varepsilon}g_{11}(\grave{x})g_{21}^T(\grave{x})\\ \frac{1}{\varepsilon}g_{21}(\grave{x})g_{11}^T(\grave{x}) & \frac{1}{\varepsilon^2}g_{21}(\grave{x})g_{21}^T(\grave{x})\end{bmatrix}\begin{bmatrix}\grave{V}_{\grave{x}_1}^T(\grave{x},y)\\ \grave{V}_{\grave{x}_2}^T(\grave{x},y)\end{bmatrix}\\ &-\tfrac{3}{2}(y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2))^T(y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2))=0,\quad \grave{V}(0,0)=0,\end{aligned}\tag{4.47}$$

together with the side-conditions

$$\grave{V}_{\grave{x}_1}(\grave{x},y)\grave{L}_1(\grave{x},y) = -(y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2))^T \tag{4.48}$$

$$V_{\grave{x}_2}(\grave{x},y)\grave{L}_2(\grave{x},y) = -\varepsilon(y-h_{21}(\grave{x}_1)-h_{22}(\grave{x}_2))^T. \tag{4.49}$$

Then, the filter $\mathbf{F}^a_{3ag}$ with

$$\grave{w}^\star=\frac{1}{\gamma^2}[g_{11}^T(\grave{x})\grave{V}_{\grave{x}_1}^T(\grave{x},y)+\frac{1}{\varepsilon}g_{21}^T(\grave{x})\grave{V}_{\grave{x}_2}^T(\grave{x},y)]$$

solves the $\mathcal{H}_\infty$ local filtering problem for the system.

Proof: Proof follows along the same lines as Proposition 4.1.1. $\square$

The result of Theorem 4.1.2 can similarly be specialized to the linear system $\mathbf{P}^l_{sp}$. Also, based on a local linearization of the system, bounds on ε and γ can be obtained similar to the result of Corollary 4.1.2.

Remark 4.1.5. *Also, comparing the accuracy of the filters $\mathbf{F}^a_{1c}$, $\mathbf{F}^a_{2c}$, $\mathbf{F}^a_{3ga}$, we see that the order of the accuracy is $\mathbf{F}^a_{2c} \succeq \mathbf{F}^a_{1c} \succeq \mathbf{F}^a_{3ag}$ by virtue of the decomposition, where the relational operator " $\succeq$" implies better.*

To obtain the limiting filter (4.46) as $\varepsilon \downarrow 0$, we must obtain the reduced-order model of the system (4.2), since $\grave{w}^\star$ is unbounded as $\varepsilon \downarrow 0$. Using Assumption 4.1.1, i.e., the equation

$$0 = f_2(x_1, x_2) + \tilde{g}_{21}(x_1, x_2)w \tag{4.50}$$

has $k \geq 1$ isolated roots, we can denote any one of these roots by

$$\bar{x}_2 = p(x_1, w), \tag{4.51}$$

for some smooth function $p : \mathcal{X} \times \mathcal{W} \to \mathcal{X}$. Then, we have the reduced-order system

$$\mathbf{P}^a_{spr} \; : \; \left\{ \begin{array}{rcl} \dot{x}_1 & = & f_1(x_1, \bar{x}_2) + g_{11}(x_1, \bar{x}_2)w; \;\; x_1(t_0) = x_{10} \\ y & = & h_{21}(x_1) + h_{22}(\bar{x}_2) + k_{21}(x_1, \bar{x}_2)w, \end{array} \right. \tag{4.52}$$

and the corresponding reduced-order filter is given by

$$\mathbf{F}^a_{3agr} \; : \; \left\{ \begin{array}{rcl} \dot{\acute{x}}_1 & = & f_1(\acute{x}_1, p(\acute{x}_1, \acute{w}^\star)) + g_{11}(\acute{x}_1, p(\acute{x}_1, \acute{w}^\star))\acute{w}^\star + \\ & & \acute{L}_1(\acute{x}, y)(y - h_{21}(\acute{x}_1) + h_{22}(p(\acute{x}_1, \acute{w}^\star)); \quad \acute{x}_1(t_0) = 0 \\ \acute{z} & = & y - h_{21}(\acute{x}_1) + h_{22}(p(\acute{x}_1, \acute{w}^\star)), \end{array} \right. \tag{4.53}$$

where all the variables have their corresponding previous meanings and dimensions, while

$$\acute{w}^{\star} = \frac{1}{\gamma^2} g_{11}^T(\acute{x}) \acute{V}_{\acute{x}_1}^T(\acute{x}, y),$$

$$\acute{V}_{\acute{x}_1}(\acute{x}, y)\acute{L}_1(\acute{x}, y) = -(y - h_{21}(\grave{x}_1) - h_{22}(p(\acute{x}_1, \acute{w}^{\star}))^T,$$

with $\acute{V}$ satisfying the following HJIE:

$$\begin{aligned}
&\acute{V}_{\acute{x}_1}(\acute{x}, y) f_1(\acute{x}_1, p(\acute{x}_1, \acute{w}^{\star})) + \acute{V}_y(\acute{x}_1, y) \nabla_{\grave{x}_1} h_{21}(\acute{x}_1) f_1(\acute{x}_1, p(\acute{x}_1, \acute{w}^{\star}))+ \\
&\tfrac{1}{2\gamma^2}[\acute{V}_{\acute{x}_1}(\acute{x}_1, y) + 2\acute{V}_y(\acute{x}_1, y)\nabla_{\grave{x}_1} h_{21}(\acute{x}_1)] g_{11}(\acute{x}, p(\acute{x}_1, \acute{w}^{\star})) g_{11}^T(\acute{x}, p(\acute{x}_1, \acute{w}^{\star})) \acute{V}_{\acute{x}_1}^T(\acute{x}_1, y) - \\
&\tfrac{1}{2}(y - h_{21}(\acute{x}_1) - h_{22}(p(\acute{x}_1, \acute{w}^{\star}))^T (y - h_{21}(\acute{x}_1) - h_{22}(p(\acute{x}_1, \acute{w}^{\star})) = 0, \;\; \acute{V}(0, 0) = 0. \quad (4.54)
\end{aligned}$$

4.1.4 Push-Pull Configuration

Finally, in this subsection, we present a *"push-pull"* configuration for the aggregate filter presented in the above section. Since the dynamics of the second subsystem is fast, we can afford to reduce the gain of the filter for this subsystem to avoid instability, while for the slow subsystem, we can afford to increase the gain. Therefore, we consider the following filter configuration

$$\mathbf{F}_{4ag}^a : \begin{cases} \dot{\check{x}}_1 &= f_1(\check{x}) + g_{11}(\check{x}_1, \check{x}_2)\check{w}^{\star} + (\check{L}_1 + \check{L}_2)(\check{x}, y)(y - h_{21}(\check{x}_1) + h_{22}(\check{x}_2)); \\ & \check{x}_1(t_0) = 0 \\ \varepsilon\dot{\check{x}}_2 &= f_2(\check{x}) + g_{11}(\check{x}_1, \check{x}_2)\check{w}^{\star} + (\check{L}_1 - \check{L}_2)(\check{x}, y)(y - h_{21}(\check{x}_1) + h_{22}(\check{x}_2)); \\ & \check{x}_2(t_0) = 0 \\ \check{z} &= y - h_{21}(\check{x}_1) + h_{22}(\check{x}_2), \end{cases} \quad (4.55)$$

where $\check{x} \in \mathcal{X}$ is the filter state, $\check{L}_1 \in \Re^{n_1 \times m}$, $\check{L}_2 \in \Re^{n_2 \times m}$ are the filter gains, while all the other variables have their corresponding previous meanings and dimensions.

Again, going through similar manipulations as in Proposition 4.1.1 we can arrive at the following result.

Proposition 4.1.2. *Consider the nonlinear system (4.2) and the $\mathcal{H}_\infty$ local filtering problem for this system. Suppose the plant $\mathbf{P^a_{sp}}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and zero-input observable for all $\varepsilon \in [0, \varepsilon^\star)$. Further, suppose for some $\gamma > 0$ and $\varepsilon \in [0, \varepsilon^\star)$, there exist a C^1 positive-semidefinite function $\check{V} : \check{N} \times \check{\Upsilon} \to \Re_+$, locally defined in a neighborhood $\check{N} \times \check{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\check{x}_1, \check{x}_2, y) = (0, 0, 0)$, and matrix functions $\check{L}_1 \in \Re^{n_1 \times m}$, $\check{L}_2 :\in \Re^{n_2 \times m}$, satisfying the HJIE (4.47) together with the side-conditions*

$$
\begin{aligned}
(\check{V}_{\check{x}_1} + \check{V}_{\check{x}_2})(\check{x}, y)\check{L}_1(\check{x}, y) &= -(y - h_{21}(\check{x}_1) - h_{22}(\check{x}_2))^T && (4.56)\\
(\check{V}_{\check{x}_1} - \check{V}_{\check{x}_2})(\check{x}, y)\check{L}_2(\check{x}, y) &= -\varepsilon(y - h_{21}(\check{x}_1) - h_{22}(\check{x}_2))^T. && (4.57)
\end{aligned}
$$

Then, the filter $\mathbf{F}^a_{4ag}$ solves the $\mathcal{H}_\infty$ local filtering problem for the system.

In the next section, we consider some examples.

4.1.5 Examples

Consider the following singularly-perturbed nonlinear system

$$
\begin{aligned}
\dot{x}_1 &= -x_1^3 + x_2\\
\varepsilon\dot{x}_2 &= -x_1 - x_2 + w\\
y &= x_1 + x_2 + w,
\end{aligned}
$$

where $w \in \mathcal{L}_2[0, \infty)$, $\varepsilon \geq 0$. We construct the aggregate filter $\mathbf{F}^a_{3ag}$ presented in the previous section for the above system. It can be checked that the system is locally observable, and the function $\grave{V}(\grave{x}) = \frac{1}{2}(\grave{x}_1^2 + \varepsilon\grave{x}_2^2)$, solves the inequality form of the HJIE (4.47) corresponding to the system. Subsequently, we calculate the gains of the aggregate filter as

$$
\grave{L}_1(\grave{x}, y) = -\frac{(y - \grave{x}_1 - \grave{x}_2)}{\grave{x}_1}, \quad \grave{L}_2(\grave{x}, y) = -\frac{\varepsilon(y - \grave{x}_1 - \grave{x}_2)}{\grave{x}_2}, \qquad (4.58)
$$

where $\grave{L}_1(\grave{x}, y)$, $\grave{L}_2(\grave{x}, y)$ are set equal to zero if $|\grave{x}_1| < \epsilon$ (small), $|\grave{x}_2| < \epsilon$ (small) respectively to avoid the singularity at $\grave{x} = 0$.

Similarly, we can construct the push-pull filter gains for the above system as

$$\check{L}_1(\check{x}, y) = -\frac{(y - \check{x}_1 - x_2^{\flat})}{\check{x}_1 + \check{x}_2}, \quad \check{L}_2(\check{x}, y) = -\frac{\varepsilon(y - \check{x}_1 - \check{x}_2)}{\check{x}_1 + \check{x}_2}. \tag{4.59}$$

4.2 $\mathcal{H}_\infty$ Filtering for Discrete-time Systems

In this section, we discuss the corresponding $\mathcal{H}_\infty$ filtering results for discrete-time singularly-perturbed affine nonlinear systems. We similarly discuss decomposition, aggregate and reduced-order filters.

4.2.1 Problem Definition and Preliminaries

The general set-up for studying discrete-time $\mathcal{H}_\infty$ filtering problems is shown in Fig. 4.2, where $\mathbf{P}_k$ is the plant, while $\mathbf{F}_k$ is the filter. The noise signal $w \in \mathcal{P}'$ is in general a bounded power signal (e.g. a Gaussian white-noise signal) which belongs to the set $\mathcal{P}'$ of bounded spectral signals, and similarly $\tilde{z} \in \mathcal{P}'$, is also a bounded power signal or ℓ_2 signal. Thus, the induced norm from w to $\tilde{z}$ (the penalty variable to be defined later) is the ℓ_∞-norm of the interconnected system $\mathbf{F}_k \circ \mathbf{P}_k$, i.e., i.e.,

$$\|\mathbf{F}_k \circ \mathbf{P}_k\|_{\ell_\infty} \triangleq \sup_{0 \neq w \in \mathcal{S}'} \frac{\|\tilde{z}\|_{\mathcal{P}'}}{\|w\|_{\mathcal{P}'}}, \tag{4.60}$$

where

$$\mathcal{P}' \triangleq \{w : \ w \in \ell_\infty, \ R_{ww}(k), \ S_{ww}(j\omega) \text{ exist for all } k \text{ and all } \omega \text{ resp.}, \|w\|_{\mathcal{P}'} < \infty\}$$

$$\|z\|_{\mathcal{P}'}^2 \triangleq \lim_{K \to \infty} \frac{1}{2K} \sum_{k=-K}^{K} \|z_k\|^2,$$

and R_{ww}, $S_{ww}(j\omega)$ are the autocorrelation and power spectral density matrices of w. Notice also that, $\|(.)\|_{\mathcal{P}'}$ is a seminorm. In addition, if the plant is stable, we replace the induced ℓ_∞-norm above by the equivalent $\mathcal{H}_\infty$ subspace norms.

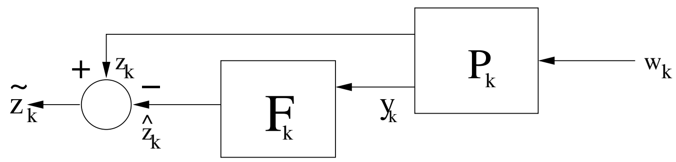

Figure 4.2 Set-up for discrete-time $\mathcal{H}_\infty$ filtering

At the outset, we consider the following singularly-perturbed affine nonlinear causal discrete-time state-space model of the plant which is defined on $\mathcal{X} \subseteq \Re^{n_1+n_2}$ with zero control input:

$$\mathbf{P}^{da}_{sp} : \begin{cases} x_{1,k+1} &= f_1(x_{1,k}, x_{2,k}, \varepsilon) + g_{11}(x_{1,k}, x_{2,k})w_k; \quad x_1(k_0, \varepsilon) = x^{10} \\ \varepsilon x_{2,k+1} &= f_2(x_{1,k}, x_{2,k}) + g_{21}(x_{1,k}, x_{2,k})w_k; \quad x_2(k_0, \varepsilon) = x^{20} \\ y_k &= h_{21}(x_{1,k}) + h_{22}(x_{2,k}) + k_{21}(x_{1,k}, x_{2,k})w_k, \end{cases} \tag{4.61}$$

where $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathcal{X}$ is the state vector with x_1 the slow state which is n_1-dimensional and x_2 the fast, which is n_2-dimensional; $w \in \mathcal{W} \subseteq \Re^r$ is an unknown disturbance (or noise) signal, which belongs to the set $\mathcal{W} \subset \ell_2[k_0, \infty) \subset \mathcal{P}'$ of admissible exogenous inputs; $y \in \mathcal{Y} \subset \Re^m$ is the measured output (or observation) of the system, and belongs to $\mathcal{Y}$, the set of admissible measured-outputs; while ε is a small perturbation parameter.

The functions $f_1 : \mathcal{X} \to \Re^{n_1}$, $\mathcal{X} \subset \Re^{n_1+n_2}$, $f_2 : \mathcal{X} \times \Re \to \Re^{n_2}$, $g_{11} : \mathcal{X} \to \mathcal{M}^{n_1 \times r}(\mathcal{X})$, $g_{21} : \mathcal{X} \to \mathcal{M}^{n_2 \times r}(\mathcal{X})$, where $\mathcal{M}^{i \times j}$ is the ring of $i \times j$ smooth matrices over $\mathcal{X}$, $h_{21}, h_{22} : \mathcal{X} \to \Re^m$, and $k_{21} : \mathcal{X} \to \mathcal{M}^{m \times r}(\mathcal{X})$ are real C^∞ functions of x. More specifically, f_2 is of the form $f_2(x_{1,k}, x_{2,k}, \varepsilon) = (\varepsilon x_{2,k} + \bar{f}_2(x_{1,k}, x_{2.k})$ for some smooth function $\bar{f}_2 : \mathcal{X} \to \Re^{n_2}$. Furthermore, we assume without any loss of generality that the system (4.61) has an isolated equilibrium-point at $(x_1^T, x_2^T) = (0, 0)$ such that $f_1(0, 0) = 0$, $f_2(0, 0) = 0$, $h_{21}(0, 0) = h_{22}(0, 0) = 0$. We also assume that there exists a unique solution $x(k, k_0, x_0, w, \varepsilon)$ $\forall k \in \mathbf{Z}$ for the system, for all initial conditions $x(k_0) \stackrel{\Delta}{=} x^0 = (x^{10^T},\ x^{20^T})^T$, for all $w \in \mathcal{W}$, and all $\varepsilon \in \Re$.

The suboptimal $\mathcal{H}_\infty$ local filtering/state estimation problem is defined as follows.

Definition 4.2.1. *(Sub-optimal $\mathcal{H}_\infty$ Local State Estimation (Filtering) Problem). Find a filter, $\mathbf{F}_k$, for estimating the state x_k or a function of it, $z_k = h_1(x_k)$, from observations*

$\mathsf{Y}_k \triangleq \{y_i : i \leq k\}$ *of* y_i *up to time* k*, to obtain the estimate*

$$\hat{x}_k = \mathbf{F}_k(\mathsf{Y}_k),$$

such that, the $\mathcal{H}_\infty$*-norm from the input* $w \in \mathcal{W}$ *to some suitable penalty function* z *is locally rendered less than or equal to a given number* γ *for all initial conditions* $x^0 \in \mathcal{O} \subset \mathcal{X}$*, for all* $w \in \mathcal{W} \subset \ell_2([k_0, \infty), \Re^r)$*. Moreover, if the filter solves the problem for all* $x^0 \in \mathcal{X}$*, we say the problem is solved globally.*

In the above definition, the condition that the $\mathcal{H}_\infty$-norm is less than or equal to γ, is more correctly referred to as the ℓ_2-gain condition

$$\sum_{k_0}^{\infty} \|z_k\|^2 \leq \gamma^2 \sum_{k_0}^{\infty} \|w_k\|^2, \quad x^0 \in \mathcal{O} \subset \mathcal{X}, \ \forall w \in \mathcal{W}. \tag{4.62}$$

We shall adopt the following notion of local observability.

Definition 4.2.2. *For the nonlinear system* $\mathbf{P}^a_{sp}$*, we say that, it is locally zero-input observable, if for all states* $x_1, \ x_2 \in U \subset \mathcal{X}$ *and input* $w(.) = 0$

$$y(k; x_1, w) \equiv y(k; x_2, w) \Longrightarrow x_1 = x_2,$$

where $y(., x_i, w), i = 1, 2$ *is the output of the system with the initial condition* $x_{k_0} = x_i$*. Moreover, the system is said to be zero-input observable if it is locally zero-input observable at each* $x^0 \in \mathcal{X}$ *or* $U = \mathcal{X}$*.*

4.2.2 Solution to the Discrete-time $\mathcal{H}_\infty$ Filtering Problem Using Decomposition Filters

In this section, we present a decomposition approach to the $\mathcal{H}_\infty$ estimation problem defined in the previous section, while in the next section, we present an aggregate approach.

We construct two time-scale filters corresponding to the decomposition of the system into a *"fast"* and *"slow"* subsystems. As in the linear case (Aganovic, 1996), (Chang, 1972), (Kim 2002), (Lim, 1996), (Sadjadi, 1990), we first assume that there exists locally a smooth invertible coordinate transformation (a diffeomorphism) $\varphi : x \mapsto \xi$, i.e.,

$$\xi_1 = \varphi_1(x, \varepsilon), \quad \varphi_1(0, \varepsilon) = 0, \quad \xi_2 = \varphi_2(x, \varepsilon), \quad \varphi_2(0, \varepsilon) = 0, \;\; \xi_1 \in \Re^{n_1}, \xi_2 \in \Re^{n_2}, \tag{4.63}$$

such that the system (4.61) is locally decomposed into the form

$$\tilde{\mathbf{P}}^{\mathbf{da}}_{sp} : \begin{cases} \xi_{1,k+1} &= \tilde{f}_1(\xi_{1,k}, \varepsilon) + \tilde{g}_{11}(\xi_k, \varepsilon)w_k, \quad \xi_1(k_0) = \varphi_1(x^0, \varepsilon) \\ \varepsilon\xi_{2,k+1} &= \tilde{f}_2(\xi_{2,k}, \varepsilon) + \tilde{g}_{21}(\xi_k, \varepsilon)w_k; \quad \xi_2(k_0) = \varphi_2(x^0, \varepsilon) \\ y_k &= \tilde{h}_{21}(\xi_{1,k}, \xi_{2,k}, \varepsilon) + \tilde{h}_{22}(\xi_{1,k}, \xi_{2,k}, \varepsilon) + \tilde{k}_{21}(\xi_k, \varepsilon)w. \end{cases} \tag{4.64}$$

Remark 4.2.1. *It is virtually impossible to find a coordinate transformation such that* $\tilde{h}_{2j} = \tilde{h}_{2j}(\xi_j), j = 1, 2$. *Thus, we have made the more practical assumption that* $\tilde{h}_{2j} = \tilde{h}_{2j}(\xi_1, \xi_2), j = 1, 2$.

Necessary conditions that such a transformation must satisfy are given in (Aliyu, 2011a). The filter is then designed based on this transformed model as follows

$$\mathbf{F}^{da}_{1c} : \begin{cases} \hat{\xi}_{1,k+1} &= \tilde{f}_1(\hat{\xi}_{1,k}, \varepsilon) + \tilde{g}_{11}(\hat{\xi}_k, \varepsilon)w^\star_k + L_1(\hat{\xi}_k, y_k, \varepsilon)[y_k - \tilde{h}_{21}(\hat{\xi}_k, \varepsilon) - \tilde{h}_{22}(\hat{\xi}_k, \varepsilon)]; \\ & \hat{\xi}_1(k_0, \varepsilon) = 0 \\ \varepsilon\hat{\xi}_{2,k+1} &= \tilde{f}_2(\hat{\xi}_{2,k}, \varepsilon) + \tilde{g}_{21}(\hat{\xi}_k, \varepsilon)w^\star_k + L_2(\hat{\xi}_k, y_k, \varepsilon)[y_k - \tilde{h}_{21}(\hat{\xi}_k, \varepsilon) - \tilde{h}_{22}(\hat{\xi}_k, \varepsilon)]; \\ & \hat{\xi}_2(k_0, \varepsilon) = 0, \end{cases} \tag{4.65}$$

where $\hat{\xi} \in \mathcal{X}$ is the filter state, $L_1 \in \Re^{n_1 \times m}$, $L_2 \in \Re^{n_2 \times m}$ are the filter gains, and $w^\star$ is the worst-case noise, while all the other variables have their corresponding previous meanings and dimensions. We can then define the penalty variable or estimation error at each instant k as

$$\tilde{z}_k = y_k - \tilde{h}_{21}(\hat{\xi}_k) - \tilde{h}_{22}(\hat{\xi}_k). \tag{4.66}$$

The problem can then be formulated as a dynamic optimization problem with the following cost functional

$$\begin{aligned}\min_{L_1\in\Re^{n_1\times m},L_2\in\Re^{n_2\times m}} \sup_{w\in\mathcal{W}} J_1(L_1,L_2,w) &= \frac{1}{2}\sum_{k=k_0}^{\infty}\left\{\|\tilde{z}_k\|^2-\gamma^2\|w_k\|^2\right\},\ s.t.\ (4.65),\\ &\quad \text{and with } w=0,\ \lim_{k\to\infty}\{\hat{\xi}_k-\xi_k\}=0. \quad (4.67)\end{aligned}$$

To solve the problem, we form the Hamiltonian function $H:\mathcal{X}\times\mathcal{W}\times\mathcal{Y}\times\Re^{n_1\times m}\times\Re^{n_2\times m}\times\Re\to\Re$:

$$\begin{aligned}H(\hat{\xi},w,y,L_1,L_2,V,\varepsilon) &= V\Big(\tilde{f}_1(\hat{\xi}_1,\varepsilon)+\tilde{g}_{11}(\hat{\xi},\varepsilon)w+L_1(\hat{\xi},y,\varepsilon)(y-\tilde{h}_{21}(\hat{\xi}_1,\varepsilon)-\\ &\quad h_{22}(\hat{\xi}_2,\varepsilon)),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon)+\tilde{g}_{21}(\hat{\xi},\varepsilon)w+\frac{1}{\varepsilon}L_2(\hat{\xi},y,\varepsilon)(y-\\ &\quad \tilde{h}_{21}(\hat{\xi},\varepsilon)-\tilde{h}_{22}(\hat{\xi},\varepsilon)),y\Big)-V(\hat{\xi},y_{k-1})+\\ &\quad \frac{1}{2}(\|\tilde{z}\|^2-\gamma^2\|w\|^2) \quad (4.68)\end{aligned}$$

for some C^1 positive-definite function $V:\mathcal{X}\times\mathcal{Y}\to\Re_+$ and where $\hat{\xi}_1=\hat{\xi}_{1,k}$, $\hat{\xi}_2=\hat{\xi}_{2,k}$ $y=y_k$, $z=\{z_k\}$, $w=\{w_k\}$. We then determine the worst-case noise $w^\star$ and the optimal gains $\hat{L}_1^\star$ and $\hat{L}_2^\star$ by maximizing and minimizing H with respect to w and L_1, L_2 respectively in the above expression (4.68), as

$$\begin{aligned}w^\star &= \arg\sup\nolimits_w H(\hat{\xi},w,y,L_1,L_2,V,\varepsilon) \quad (4.69)\\ [L_1^\star,L_2^\star] &= \arg\min_{L_1,L_2} H(\hat{\xi},w^\star,y,L_1,L_2,V,\varepsilon). \quad (4.70)\end{aligned}$$

However, because the Hamiltonian function (4.68) is not a linear or quadratic function of w and L_1, L_2, only implicit solutions may be obtained (Aliyu, 2011a). Thus, the only way to obtain an explicit solution is to use an approximate scheme. In (Aliyu, 2011a) we have used a second-order Taylor series approximationn of the Hamiltonian about $(\tilde{f}_1(\hat{\xi}_1),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2),y)$ in the direction of the state vectors $(\hat{\xi}_1,\hat{\xi}_2)$. It is believed that, this would capture most, if not all, of the system dynamics. However, for the $\mathcal{H}_\infty$ problem at hand, such an approximation becomes too messy and the solution becomes more involved. Therefore, instead we would rather use a first-order Taylor approximation which is given by

$$\begin{aligned}
\widehat{H}(\hat{\xi},\hat{w},y,\hat{L}_1,\hat{L}_2,\hat{V},\varepsilon) \;=\;& \hat{V}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)-\hat{V}(\hat{\xi},y_{k-1})+\\
& \hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)[\tilde{g}_{11}(\hat{\xi},\varepsilon)\hat{w}+\\
& \hat{L}_1(\hat{\xi},y,\varepsilon)(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon)]+\\
& \frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_{2,\varepsilon}}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)[\tilde{g}_{21}(\hat{\xi},\varepsilon)\hat{w}+\\
& \hat{L}_2(\hat{\xi},y,\varepsilon)(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon)]+\\
& \frac{1}{2}(\|\tilde{z}\|^2-\gamma^2\|\hat{w}\|^2)+O(\|\hat{\xi}\|^2)
\end{aligned}\tag{4.71}$$

where $\hat{V}$, $\hat{w}$, $\hat{L}_1$, $\hat{L}_2$ are the corresponding approximate functions, and $\hat{V}_{\hat{\xi}_1}$, $\hat{V}_{\hat{\xi}_2}$ are the row vectors of first-partial derivatives of $\hat{V}$ with respect to $\hat{\xi}_1$, $\hat{\xi}_2$ respectively. We can now obtain $w^\star$ as

$$\hat{w}^\star=\frac{1}{\gamma^2}[\tilde{g}_{11}^T(\hat{\xi},\varepsilon)\hat{V}_{\hat{\xi}_1}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)+\frac{1}{\varepsilon}\tilde{g}_{21}^T(\hat{\xi},\varepsilon)\hat{V}_{\hat{\xi}_2}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y))\tag{4.72}$$

Then, substituting $\hat{w}=\hat{w}^\star$ in (4.71), we have

$$\begin{aligned}
&\widehat{H}(\hat{\xi},\hat{w}^\star,y,\hat{L}_1,\hat{L}_2,\hat{V},\varepsilon)\approx\hat{V}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)-\hat{V}(\hat{\xi},y_{k-1})+\\
&\quad\frac{1}{2\gamma^2}\Big[\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\tilde{g}_{11}(\hat{\xi},\varepsilon)\tilde{g}_{11}^T(\hat{\xi},\varepsilon)\hat{V}_{\hat{\xi}_1}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)+\\
&\quad\frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\tilde{g}_{11}(\hat{\xi},\varepsilon)\tilde{g}_{21}^T(\hat{\xi},\varepsilon)\hat{V}_{\hat{\xi}_{2,\varepsilon}}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\Big]+\\
&\quad\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\hat{L}_1(\hat{\xi},y,\varepsilon)(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon))+\\
&\quad\frac{1}{2\gamma^2}\Big[\frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2),y)\tilde{g}_{21}(\hat{\xi},\varepsilon)\tilde{g}_{11}^T(\hat{\xi},\varepsilon)\hat{V}_{\hat{\xi}_1}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)+\\
&\quad\frac{1}{\varepsilon^2}\hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\tilde{g}_{21}(\hat{\xi})\tilde{g}_{21}^T(\hat{\xi})\hat{V}_{\hat{\xi}_2}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\Big]+\\
&\quad\frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\hat{L}_2(\hat{\xi},y,\varepsilon)(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon))+\frac{1}{2}\|\tilde{z}\|^2.
\end{aligned}\tag{4.73}$$

Completing the squares now for $\hat{L}_1(\hat{\xi},y)$ and $\hat{L}_2(\hat{\xi},y)$ in (4.73), we get

$$\widehat{H}(\hat{\xi},\hat{w}^\star,y,\hat{L}_1,\hat{L}_2,\hat{V},\varepsilon)\approx\hat{V}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)-\hat{V}(\hat{\xi},y_{k-1})$$

$$
\begin{aligned}
&+\frac{1}{2\gamma^2}\Big[\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\tilde{g}_{11}(\hat{\xi},\varepsilon)\tilde{g}_{11}^T(\hat{\xi},\varepsilon)\hat{V}_{\hat{\xi}_1}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\\
&+\frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\tilde{g}_{11}(\hat{\xi},\varepsilon)\tilde{g}_{21}^T(\hat{\xi},\varepsilon)\hat{V}_{\hat{\xi}_2}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\Big]\\
&+\frac{1}{2}\left\|\hat{L}_1^T(\hat{\xi},y)\hat{V}_{\hat{\xi}_1}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)+(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon))\right\|^2+\\
&\frac{1}{2\gamma^2}\Big[\frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\tilde{g}_{21}(\hat{\xi},\varepsilon)\tilde{g}_{11}^T(\hat{\xi},\varepsilon)\hat{V}_{\hat{\xi}_1}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\\
&+\frac{1}{\varepsilon^2}\hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2),y)\tilde{g}_{21}(\hat{\xi})\tilde{g}_{21}^T(\hat{\xi})\hat{V}_{\hat{\xi}_2}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\Big]-\\
&\frac{1}{2}\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2),y)\hat{L}_1(\hat{\xi},y,\varepsilon)\hat{L}_1^T(\hat{\xi},y,\varepsilon)\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)-\\
&\frac{1}{2\varepsilon^2}\hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\hat{L}_2(\hat{\xi},y,\varepsilon)\hat{L}_2^T(\hat{\xi},y,\varepsilon)\hat{V}_{\hat{\xi}_2}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\\
&+\frac{1}{2}\left\|\frac{1}{\varepsilon}\hat{L}_2^T(\hat{\xi},y,\varepsilon)\hat{V}_{\hat{\xi}_2}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)+(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon))\right\|^2-\\
&\frac{1}{2}\|z\|^2. \qquad (4.74)
\end{aligned}
$$

Hence, setting the optimal gains as

$$
\begin{aligned}
\hat{V}_{\hat{\xi}_1,\varepsilon}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\hat{L}_1^\star(\hat{\xi},y,\varepsilon) &= -(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon))^T \qquad (4.75)\\
\hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\hat{L}_2^\star(\hat{\xi},y,\varepsilon) &= -\varepsilon(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon))^T \qquad (4.76)
\end{aligned}
$$

minimizes the Hamiltonian $\widehat{H}(.,.,\hat{L}_1,\hat{L}_2,.,.)$ and guarantees that the saddle-point condition (Basar, 1982)

$$
\widehat{H}(.,\hat{w}^\star,\hat{L}_1^\star,\hat{L}_2^\star,.,.)\le\widehat{H}(.,\hat{w}^\star,\hat{L}_1,\hat{L}_2,.,.)\quad\forall\hat{L}_1\in\Re^{n_1\times m},\hat{L}_2\in\Re^{n_2\times m} \qquad (4.77)
$$

is satisfied. Finally, substituting the above optimal gains in (4.71) and setting

$$
\widehat{H}(\hat{\xi},w^\star,y,\hat{L}_1^\star,\hat{L}_2^\star,\hat{V},\varepsilon)=0,
$$

results in the following discrete Hamilton-Jacobi-Isaacs equation (DHJIE):

$$
\hat{V}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)-\hat{V}(\hat{\xi},y_{k-1})+
$$

$$\frac{1}{2\gamma^2}[\;\hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\tfrac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\quad \hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\tfrac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\;]\times$$
$$\left[\begin{array}{cc} \tilde{g}_{11}(\hat{\xi})\tilde{g}_{11}^T(\hat{\xi},\varepsilon) & \frac{1}{\varepsilon}\tilde{g}_{11}(\hat{\xi},\varepsilon)\tilde{g}_{21}^T(\hat{\xi},\varepsilon) \\ \frac{1}{\varepsilon}\tilde{g}_{21}(\hat{\xi},\varepsilon)\tilde{g}_{11}^T(\hat{\xi},\varepsilon) & \frac{1}{\varepsilon^2}\tilde{g}_{21}(\hat{\xi},\varepsilon)\tilde{g}_{21}^T(\hat{\xi},\varepsilon)\end{array}\right]\left[\begin{array}{c}\hat{V}_{\hat{\xi}_1}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y) \\ \hat{V}_{\hat{\xi}_2}^T(\tilde{f}_1(\hat{\xi}_1,\varepsilon),\frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2,\varepsilon),y)\end{array}\right] -$$
$$\frac{3}{2}(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon))^T(y-\tilde{h}_{21}(\hat{\xi},\varepsilon)-h_{22}(\hat{\xi},\varepsilon)) = 0 \quad \hat{V}(0,0,0)=0. \tag{4.78}$$

We then have the following result.

Proposition 4.2.1. *Consider the nonlinear discrete system (4.61) and the $\mathcal{H}_\infty$-filtering problem for this system. Suppose the plant $\mathbf{P}_{\mathbf{sp}}^{\mathbf{da}}$ is locally asymptotically stable about the equilibrium-point $x=0$ and zero-input observable. Further, suppose there exist a local diffeomorphism φ that transforms the system to the partially decoupled form (4.64), a C^1 positive-semidefinite function $\hat{V}:\hat{N}\times\hat{\Upsilon}\to\Re_+$ locally defined in a neighborhood $\hat{N}\times\hat{\Upsilon}\subset\mathcal{X}\times\mathcal{Y}$ of the origin $(\hat{\xi},y)=(0,0)$, and matrix functions $\hat{L}_i:\hat{N}\times\hat{\Upsilon}\to\Re^{n_i\times m}$, $i=1,2$, satisfying the DHJIE (4.78) together with the side-conditions (4.75), (4.76) for some $\gamma>0$. Then, the filter $\mathbf{F}_{1c}^{da}$ solves the $\mathcal{H}_\infty$ filtering problem for the system locally in $\hat{N}$.*

Proof: The optimality of the filter gains $\hat{L}_1^\star$, $\hat{L}_2^\star$ has already been shown above. It remains to show that the sadle-point conditions (Basar, 1982)

$$\begin{aligned}&\widehat{H}(.,\hat{w},\hat{L}_1^\star,\hat{L}_2^\star,.,.)\le\widehat{H}(.,\hat{w}^\star,\hat{L}_1^\star,\hat{L}_2^\star,.,.)\le\widehat{H}(.,\hat{w}^\star,\hat{L}_1,\hat{L}_2,.,.),\\ &\qquad\forall\hat{L}_1\in\Re^{n_1\times m},\hat{L}_2\in\Re^{n_2\times m},\forall w\in\ell_2[k_0,\infty),\end{aligned} \tag{4.79}$$

and the ℓ_2-gain condition (4.62) hold for all $w\in\mathcal{W}$. In addition, it is required aklso to show that there is asymptotic convergence of the estimation error vector.

Now, the right-hand-side of the above inequality (4.79) has already been shown. It remains to show that the left hand side also holds. Accordingly, it can be shown from (4.71), (4.78) that

$$\widehat{H}(\hat{\xi},\hat{w},\hat{L}_1^\star,\hat{L}_2^\star,\hat{V},\varepsilon)=\widehat{H}(\hat{\xi},\hat{w}^\star,\hat{L}_1^\star,\hat{L}_2^\star,\hat{V},\varepsilon)-\frac{1}{2}\gamma^2\|\hat{w}-\hat{w}^\star\|^2$$

Therefore, we also have the left-hand side of (4.79) satisfied, and the pair $(\hat{w}^\star,[\hat{L}_1^\star,L_2^\star])$ constitute a saddle-point solution to the dynamic game (4.67), (4.65).

Next, let $\hat{V} \geq 0$ be a C^1 solution of the DHJIE (4.78). Then, consider the time-variation of $\hat{V}$ along a trajectory of (4.65), with $\hat{L}_1 = \hat{L}_1^\star$, $L_2 = \hat{L}_2^\star$, and $w \in \mathcal{W}$, to get

$$\begin{aligned}
\hat{V}(\hat{\xi}_{1,k+1}, \hat{\xi}_{2,k+1}, y) \approx\ & \hat{V}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y) + \\
& \hat{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y)[\tilde{g}_{11}(\hat{\xi}, \varepsilon)\hat{w} + \hat{L}_1^\star(\hat{\xi}, y, \varepsilon)(y - \tilde{h}_{21}(\hat{\xi}, \varepsilon) - h_{22}(\hat{\xi}, \varepsilon))] + \\
& \frac{1}{\varepsilon}\hat{V}_{\hat{\xi}_2}(\tilde{f}_1(\hat{\xi}_1, \varepsilon), \frac{1}{\varepsilon}\tilde{f}_2(\hat{\xi}_2, \varepsilon), y)[\tilde{g}_{21}(\hat{\xi}, \varepsilon)w + \hat{L}_2^\star(\hat{\xi}, y, \varepsilon)(y - \tilde{h}_{21}(\hat{\xi}, \varepsilon) - h_{22}(\hat{\xi}, \varepsilon))] \\
& = \hat{V}(\hat{\xi}, y_{k-1}) - \frac{\gamma^2}{2}\|\hat{w} - \hat{w}^\star\|^2 + \frac{1}{2}(\gamma^2\|\hat{w}\|^2 - \|\tilde{z}\|^2) \quad \forall \hat{w} \in \mathcal{W} \\
& \leq \hat{V}(\hat{\xi}, y_{k-1}) + \frac{1}{2}(\gamma^2\|\hat{w}\|^2 - \|\tilde{z}\|^2) \quad \forall \hat{w} \in \mathcal{W}
\end{aligned} \tag{4.80}$$

where we have used the first-order Taylor approximation in the above, and the last inequality follows from using the DHJIE (4.78). Moreover, the last inequality is the discrete-time dissipation-inequality (Guillard, 1996), which also implies that the ℓ_2-gain inequality (4.62) is satisfied.

In addition, setting $\hat{w} = 0$ in (4.80) implies that

$$\hat{V}(\hat{\xi}_{1,k+1}, \hat{\xi}_{2,k+1}, y) - \hat{V}(\hat{\xi}_{1,k}, \hat{\xi}_{2,k}, y_{k-1}) = -\frac{1}{2}\|z_k\|^2.$$

Therefore, the filter dynamics is stable, and $\hat{V}(\hat{\xi}, y)$ is non-increasing along a trajectory of (4.65). Further, the condition that $\hat{V}(\hat{\xi}_{1,k+1}, \hat{\xi}_{2,k+1}, y) \equiv \hat{V}(\hat{\xi}_{1,k}, \hat{\xi}_{2,k}, y_{k-1})$ $\forall k \geq k_s$ (say!) implies that $z_k \equiv 0$, which further implies that $y_k = \tilde{h}_{21}(\hat{\xi}_k) + \tilde{h}_{22}(\hat{\xi}_k)$ $\forall k \geq k_s$. By the zero-input observability of the system, this implies that $\hat{\xi} = \xi$. Finally, since φ is invertible and $\varphi(0, \varepsilon) = 0$, $\hat{\xi} = \xi$ implies $\hat{x} = \varphi^{-1}(\hat{\xi}, \varepsilon) = \varphi^{-1}(\xi, \varepsilon) = x$. $\square$

Next, we consider the limiting behavior of the filter (4.65) and the corresponding DHJIE (4.78). Letting $\varepsilon \downarrow 0$, we obtain from (4.65),

$$0 = \tilde{f}_2(\hat{\xi}_{2,k}) + L_2(\hat{\xi}_k, y_k)(y_k - \tilde{h}_{21}(\hat{\xi}_k) - \tilde{h}_{22}(\hat{\xi}_k)) \quad \forall k,$$

and since $\tilde{f}_2(.)$ is asymptotically stable, we have $\hat{\xi}_2 \to 0$. Therefore, $H(., ., ., ., .)$ in (4.68) becomes

$$
\begin{aligned}
H_0(\hat{\xi}, w, y, L_1, L_2, V, 0) &= V\Big(\tilde{f}_1(\hat{\xi}_1) + \tilde{g}_{11}(\hat{\xi})w + L_1(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}_1) - h_{22}(\hat{\xi}_2)), 0, y\Big) - \\
&\quad V(\hat{\xi}, y_{k-1}) + \frac{1}{2}(\|z\|^2 - \gamma^2\|w\|^2). \qquad (4.81)
\end{aligned}
$$

A first-order Taylor approximation of this Hamiltonian about $(\tilde{f}_1(\hat{\xi}_1), 0, y)$ similarly yields

$$
\begin{aligned}
\widehat{H}_0(\hat{\xi}, \hat{w}, y, \hat{L}_{10}, \bar{V}, 0) &= \bar{V}(\tilde{f}_1(\hat{\xi}_1), 0, y) + \bar{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1), 0, y)\hat{L}_{10}^T(\hat{\xi}, y)(y - \tilde{h}_{21}(\hat{\xi}) - h_{22}(\hat{\xi})) + \\
&\quad \bar{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1), 0, y)\tilde{g}_{11}(\hat{\xi})w - \bar{V}(\hat{\xi}, y_{k-1}) + \frac{1}{2}(\|z\|^2 - \gamma^2\|\hat{w}\|^2) + \\
&\quad O(\|\hat{\xi}\|^2) \qquad (4.82)
\end{aligned}
$$

for some corresponding positive-definite function $\bar{V} : \mathcal{X} \times \mathcal{Y} \to \Re$, and gain matrix $\hat{L}_{10}$. Minimizing again this Hamiltonian, we obtain the worst-case noise $w_{10}^\star$ and optimal gain matrix $\hat{L}_{10}^\star$ given by

$$
\begin{aligned}
\hat{w}_{10}^\star &= -\tilde{g}_{11}^T(\hat{\xi})\bar{V}_{\hat{\xi}_1}^T(\tilde{f}_1(\hat{\xi}_1), 0, y), \qquad (4.83)\\
\bar{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1), 0, y)\hat{L}_{10}^\star(\hat{\xi}, y) &= -(y - \tilde{h}_{21}(\hat{\xi}) - h_{22}(\hat{\xi}))^T, \qquad (4.84)
\end{aligned}
$$

where $\bar{V}$ satisfies the reduced-order DHJIE

$$
\begin{aligned}
&\bar{V}(\tilde{f}_1(\hat{\xi}_1), 0, y) + \tfrac{1}{2\gamma^2}\bar{V}_{\hat{\xi}_1}(\tilde{f}_1(\hat{\xi}_1), 0, y)\tilde{g}_{11}(\hat{\xi})\tilde{g}_{11}^T(\hat{\xi})\bar{V}_{\hat{\xi}_1}^T(\tilde{f}_1(\hat{\xi}_1), 0, y) - \bar{V}(\hat{\xi}_1, 0, y_{k-1}) - \\
&\quad \tfrac{3}{2}(y - \tilde{h}_{21}(\hat{\xi}) - h_{22}(\hat{\xi}))^T)(y - \tilde{h}_{21}(\hat{\xi}) - h_{22}(\hat{\xi})) = 0, \quad \bar{V}(0,0,0) = 0. \qquad (4.85)
\end{aligned}
$$

The corresponding reduced-order filter is given by

$$
\bar{\mathbf{F}}_{1r}^{da} : \Big\{ \; \dot{\hat{\xi}}_1 = \tilde{f}_1(\hat{\xi}_1) + \hat{L}_{10}^\star(\hat{\xi}_1, y)(y - \tilde{h}_{21}(\hat{\xi}) - \tilde{h}_{22}(\hat{\xi})) + O(\varepsilon). \qquad (4.86)
$$

Moreover, since the gain $\hat{L}_{10}^\star$ is such that the estimation error $e_k = y_k - \tilde{h}_{21}(\hat{\xi}_k) - \tilde{h}_{22}(\hat{\xi}_k) \to 0$, and the vector-field $\tilde{f}_2(\hat{\xi}_2)$ is locally asymptotically stable, we have $\hat{L}_2^\star(\hat{\xi}_k, y_k) \to 0$ as $\varepsilon \downarrow 0$. Correspondingly, the solution $\bar{V}$ of the DHJIE (4.85) can be represented as the asymptotic limit of the solution of the DHJIE (4.78) as $\varepsilon \downarrow 0$, i.e.,

$$\hat{V}(\hat{\xi}, y) = \bar{V}(\hat{\xi}_1, y) + O(\varepsilon).$$

We can specialize the result of Proposition 4.2.1 to the following discrete-time linear singularly-perturbed system (DLSPS) (Aganovic, 1996), (Kim 2002), (Lim, 1996), (Sadjadi, 1990) in the slow coordinate:

$$\mathbf{P}^l_{dsp} : \begin{cases} x_{1,k+1} &= A_1 x_{1,k} + A_{12} x_{2,k} + B_{11} w_k; \quad x_1(k_0) = x^{10} \\ \varepsilon x_{2,k+1} &= A_{21} x_{1,k} + (\varepsilon I_{n_2} + A_2) x_{2,k} + B_{21} w_k; \quad x_2(k_0) = x^{20} \\ y_k &= C_{21} x_{1,k} + C_{22} x_{2,k} + w_k \end{cases} \tag{4.87}$$

where $A_1 \in \Re^{n_1 \times n_1}$, $A_{12} \in \Re^{n_1 \times n_2}$, $A_{21} \in \Re^{n_2 \times n_1}$, $A_2 \in \Re^{n_2 \times n_2}$, $B_{11} \in \Re^{n_1 \times s}$, and $B_{21} \in \Re^{n_2 \times s}$, while the other matrices have compatible dimensions. Then, an explicit form of the required transformation φ above is given by the Chang transformation (Chang, 1972):

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} I_{n_1} - \varepsilon \mathsf{HL} & -\varepsilon \mathsf{H} \\ \mathsf{L} & I_{n_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{4.88}$$

where the matrices L and H satisfy the equations

$$\begin{aligned} 0 &= (\varepsilon I_{n_2} + A_2)\mathsf{L} - A_{21} - \varepsilon \mathsf{L}(A_1 - A_{12}\mathsf{L}) \\ 0 &= -\mathsf{H}[(\varepsilon I_{n_2} + A_2) + \varepsilon \mathsf{L} A_{12}] + A_{12} + \varepsilon(A_1 - A_{12}\mathsf{L})\mathsf{H}. \end{aligned}$$

The system is then represented in the new coordinates by

$$\tilde{\mathbf{P}}^l_{dsp} : \begin{cases} \xi_{1,k+1} &= \tilde{A}_1 \xi_{1,k} + \tilde{B}_{11} w_k; \quad \xi_1(k_0) = \xi^{10} \\ \varepsilon \xi_{2,k+1} &= \tilde{A}_2 \xi_{2,k} + \tilde{B}_{21} w_k; \quad \xi_2(k_0) = \xi^{20} \\ y_k &= \tilde{C}_{21} \xi_{1,k} + \tilde{C}_{22} \xi_{2,k} + w_k, \end{cases} \tag{4.89}$$

where

$$\begin{aligned}
\tilde{A}_1 &= A_1 - A_{12}\mathsf{L} = A_1 - A_{12}(\varepsilon I_{n_2} + A_2)^{-1}A_{21} + O(\varepsilon)\\
\tilde{B}_{11} &= B_{11} - \varepsilon\mathsf{HL}B_{11} - \mathsf{H}B_{21} = B_{11} - A_{12}A_2^{-1}B_{21} + O(\varepsilon)\\
\tilde{A}_2 &= (\varepsilon I_{n_2} + A_2) + \varepsilon\mathsf{L}A_{12} = A_2 + O(\varepsilon)\\
\tilde{B}_{21} &= B_{21} + \varepsilon\mathsf{L}B_{11} = B_{21} + O(\varepsilon)\\
\tilde{C}_{21} &= C_{21} - C_{22}\mathsf{L} = C_{21} - C_{22}(\varepsilon I_{n_2} + A_2)^{-1}A_{21} + O(\varepsilon)\\
\tilde{C}_{22} &= C_{22} + \varepsilon(C_{21} - C_{22})\mathsf{H} = C_{22} + O(\varepsilon).
\end{aligned}$$

Adapting the filter (4.65) to the system (4.89) yields the following filter

$$\mathbf{F}^{dl}_{1c}: \left\{\begin{aligned}
\hat{\xi}_{1,k+1} &= \tilde{A}_1\hat{\xi}_{1,k} + \tilde{B}_{11}w^\star_k + \hat{L}_1(y_k - \tilde{C}_{21}\hat{\xi}_{1,k} - \tilde{C}_{22}\hat{\xi}_{2,k})\\
\varepsilon\hat{\xi}_{2,k+1} &= \tilde{A}_2\hat{\xi}_{2,k} + \tilde{B}_{21}w^\star_k + \hat{L}_2(y_k - \tilde{C}_{21}\hat{\xi}_{1,k} - \tilde{C}_{22}\hat{\xi}_{2,k}).
\end{aligned}\right. \tag{4.90}$$

Taking

$$\hat{V}(\hat{\xi}, y) = \frac{1}{2}(\hat{\xi}_1^T\hat{P}_1\hat{\xi}_1 + \hat{\xi}_2^T\hat{P}_2\hat{\xi}_2 + y^T\hat{Q}y),$$

for some symmetric positive-definite matrices $\hat{P}_1$, $\hat{P}_2$, $\hat{Q}$, the DHJIE (4.78) reduces to the following algebraic equation

$$\begin{aligned}
&(\hat{\xi}_1^T\tilde{A}_1^T\hat{P}_1\tilde{A}_1\hat{\xi}_1 + \frac{1}{\varepsilon^2}\hat{\xi}_2^T\tilde{A}_2^T\hat{P}_2\tilde{A}_2^T\hat{\xi}_2 + y^T\hat{Q}y) - (\hat{\xi}_1^T\hat{P}_1\hat{\xi}_1 + \hat{\xi}_2^T\hat{P}_2\hat{\xi}_2 + y_{k-1}^T\hat{Q}y_{k-1}) +\\
&\frac{1}{\gamma^2}\Big[\hat{\xi}_1^T\tilde{A}_1^T\hat{P}_1\tilde{B}_{11}\tilde{B}_{11}^T\hat{P}_1\tilde{A}_1\hat{\xi}_1 + \frac{1}{\varepsilon^2}\hat{\xi}_2^T\tilde{A}_2^T\hat{P}_2\tilde{B}_{21}\tilde{B}_{11}^T\hat{P}_1\tilde{A}_1\hat{\xi}_1 + \frac{1}{\varepsilon^2}\hat{\xi}_1^T\tilde{A}_1^T\hat{P}_1\tilde{B}_{11}\tilde{B}_{21}^T\hat{P}_2\tilde{A}_2\hat{\xi}_2\\
&+\frac{1}{\varepsilon^4}\hat{\xi}_2^T\tilde{A}_2^T\hat{P}_2\tilde{B}_{21}\tilde{B}_{21}^T\hat{P}_2\tilde{A}_2\hat{\xi}_2\Big] - 3(y^Ty - \hat{\xi}_1^T\tilde{C}_{21}^Ty - y^T\tilde{C}_{21}^T\hat{\xi}_1 - y^T\tilde{C}_{22}^T\hat{\xi}_1 - y^T\tilde{C}_{22}^T\hat{\xi}_2 -\\
&\hat{\xi}_2^T\tilde{C}_{22}^Ty + \hat{\xi}_1^T\tilde{C}_{21}^T\tilde{C}_{21}\hat{\xi}_1 + \hat{\xi}_1^T\tilde{C}_{21}^T\tilde{C}_{22}\hat{\xi}_2 + \hat{\xi}_2^T\tilde{C}_{22}^T\tilde{C}_{21}\hat{\xi}_1 + \hat{\xi}_2^T\tilde{C}_{22}^T\tilde{C}_{22}\hat{\xi}_2) = 0.
\end{aligned} \tag{4.91}$$

Subtracting now $\frac{1}{2}y^T\hat{R}y$ for some symmetric matrix $\hat{R} > 0$ from the left-hand side of the above equation (and absorbing $\hat{R}$ in $\hat{Q}$), we have the following matrix-inequality

$$\begin{bmatrix} \tilde{A}_1^T\hat{P}_1A_1-\hat{P}_1+\frac{1}{\gamma^2}\tilde{A}_1^T\hat{P}_1\tilde{B}_{11}\tilde{B}_{11}^T\hat{P}_1\tilde{A}_1-3\tilde{C}_{21}^T\tilde{C}_{21} & \frac{1}{\gamma^2\varepsilon^2}\tilde{A}_1^T\hat{P}_1\tilde{B}_{11}\tilde{B}_{21}^T\hat{P}_2\tilde{A}_2+3\tilde{C}_{21}^T\tilde{C}_{22} & 3\tilde{C}_{21}^T & 0 \\ \frac{1}{\gamma^2\varepsilon^2}\tilde{A}_2^T\hat{P}_2\tilde{B}_{21}\tilde{B}_{11}^T\hat{P}_1\tilde{A}_1+3\tilde{C}_{22}^T\tilde{C}_{21} & \frac{1}{\varepsilon^2}\tilde{A}_2^T\hat{P}_2\tilde{A}_2-\hat{P}_2+\frac{1}{\gamma^2\varepsilon^4}\tilde{A}_2^T\hat{P}_2\tilde{B}_{21}\tilde{B}_{21}^T\hat{P}_2\tilde{A}_2-3\tilde{C}_{22}^T\tilde{C}_{22} & 3\tilde{C}_{22}^T & 0 \\ 3\tilde{C}_{21} & 3\tilde{C}_{22} & \hat{Q}-3I & 0 \\ 0 & 0 & 0 & -\hat{Q} \end{bmatrix} \le 0, \tag{4.92}$$

while the side conditions (4.75), (4.76) reduce to the following LMIs

$$\begin{bmatrix} 0 & 0 & \frac{1}{2}(\tilde{A}_1^T\hat{P}_1\hat{L}_1-\tilde{C}_{21}^T) \\ 0 & 0 & -\frac{1}{2}\tilde{C}_{22}^T \\ \frac{1}{2}(\tilde{A}_1^T\hat{P}_1\hat{L}_1-\tilde{C}_{21}^T)^T & -\frac{1}{2}\tilde{C}_{22}^T & (1-\delta_1)I \end{bmatrix} \le 0 \tag{4.93}$$

$$\begin{bmatrix} 0 & 0 & -\frac{1}{2}\tilde{C}_{21}^T \\ 0 & 0 & \frac{1}{2\varepsilon^2}(\tilde{A}_2^T\hat{P}_2\hat{L}_2-\tilde{C}_{22}^T) \\ -\frac{1}{2}\tilde{C}_{21} & \frac{1}{2\varepsilon^2}(\tilde{A}_2^T\hat{P}_2\hat{L}_2-\tilde{C}_{22}^T)^T & (1-\delta_2)I \end{bmatrix} \le 0 \tag{4.94}$$

respectively, for some numbers δ_1, $\delta_2 \ge 1$. The above matrix inequality (4.23) can be further simplified using Schur's complements, but cannot be made linear because of the off-diagonal and coupling terms. This is primarily because the assumed transformation φ can only achieve a partial decoupling of the original system, and a complete decoupling of the states will require more stringent assumptions and conditions.

Consequently, we have the following Corollary to Proposition 4.2.1.

Corollary 4.2.1. *Consider the DLSPS (4.87) and the $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\mathbf{P}_{sp}^l$ is locally asymptotically stable about the equilibrium-point $x = 0$ and observable. Suppose further, it is transformable to the form (4.89), and there exist symmetric positive-definite matrices $\hat{P}_1 \in \Re^{n_1\times n_1}$, $\hat{P}_2 \in \Re^{n_2\times n_2}$, $\hat{Q} \in \Re^{m\times m}$, and matrices $\hat{L}_1 \in \Re^{n_1\times m}$, $\hat{L}_2 \in \Re^{n_2\times m}$, satisfying the matrix inequalities (4.92), (4.93), (4.94) for some* *numbers* $\delta_1, \delta_2 \geq 1$ *and* $\gamma > 0$. *Then, the filter* $\mathbf{F}_{1c}^{dl}$ *solves the* $\mathcal{H}_\infty$ *filtering problem for the system.*

Similarly, for the reduced-order filter (4.86) and the DHJIE (4.85), we have respectively

$$\mathbf{F}_{1r}^{dl} : \left\{ \begin{array}{rcl} \hat{\xi}_{1,k+1} & = & \tilde{A}_1\hat{\xi}_{1,k} + \hat{L}_{10}^{\star}(y_k - \tilde{C}_{21}\hat{\xi}_{1,k} - \tilde{C}_{22}\hat{\xi}_{2,k}) \end{array} \right. \tag{4.95}$$

$$\begin{bmatrix} \tilde{A}_1^T\hat{P}_{10}\tilde{A}_1 - \hat{P}_{10} - 3\tilde{C}_{21}^T\tilde{C}_{21} & \tilde{A}_1^T\hat{P}_{10}\tilde{B}_{11} & 3\tilde{C}_{21} & 0 \\ \tilde{B}_{11}^T\hat{P}_{10}\tilde{A}_1 & -\gamma^{-2}I & 0 & 0 \\ 3\tilde{C}_{21}^T & 0 & \hat{Q} - 3I & 0 \\ 0 & 0 & 0 & \hat{Q} \end{bmatrix} \leq 0 \tag{4.96}$$

$$\begin{bmatrix} 0 & 0 & \frac{1}{2}(\tilde{A}_1^T\hat{P}_{10}\hat{L}_{10} - \tilde{C}_{21}^T) \\ 0 & 0 & -\frac{1}{2}\tilde{C}_{22}^T \\ \frac{1}{2}(\tilde{A}_1^T\hat{P}_{10}\hat{L}_{10} - \tilde{C}_{21}^T)^T & -\frac{1}{2}\tilde{C}_{22}^T & (1-\delta_{10})I \end{bmatrix} \leq 0 \tag{4.97}$$

for some symmetric positive-definite matrices $\hat{P}_{10}$, $\hat{Q}_{10}$, gain matrix $\hat{L}_{10}$ and some number $\delta_{10} > 1$.

Similarly, Proposition 4.2.1 has not yet exploited the benefit of the coordinate transformation in designing the filter (4.65) for the system (4.64). We shall now design separate reduced-order filters for the decomposed subsystems which should be more efficient than the previous one. If we let $\varepsilon \downarrow 0$ in (4.64) and obtain the following reduced system model:

$$\widetilde{\mathbf{P}}_{\mathbf{r}}^{\mathbf{a}} : \left\{ \begin{array}{rcl} \xi_{1,k+1} & = & \tilde{f}_1(\xi_1) + \tilde{g}_{11}(\xi)w \\ 0 & = & \tilde{f}_2(\xi_2) + \tilde{g}_{21}(\xi)w \\ y_k & = & \tilde{h}_{21}(\xi) + \tilde{h}_{22}(\xi) + \tilde{k}_{21}(\xi)w. \end{array} \right. \tag{4.98}$$

Then, we assume the following (Khalil, 1985), (Kokotovic, 1986).

Assumption 4.2.1. *The system (4.61), (4.98) is in the "standard form", i.e., the equation*

$$0 = \tilde{f}_2(\xi_2) + \tilde{g}_{21}(\xi)w \tag{4.99}$$

has $l \geq 1$ isolated roots, we can denote any one of these solutions by

$$\bar{\xi}_2 = q(\xi_1, w) \tag{4.100}$$

for some C^1 function $q : \mathcal{X} \times \mathcal{W} \to \mathcal{X}$.

Under Assumption 4.2.1, we obtain the reduced-order slow-subsystem

$$\mathbf{P}^{\mathbf{a}}_{\mathbf{r}} : \begin{cases} \xi_{1,k+1} &= \tilde{f}_1(\xi_{1,k}) + \tilde{g}_{11}(\xi_{1,k}, q(\xi_{1,k}, w_k))w_k + O(\varepsilon) \\ y_k &= \tilde{h}_{21}(\xi_{1,k}, q(\xi_{1,k}, w_k)) + \tilde{h}_{22}(\xi_{1,k}, q(\xi_{1,k}, w_k)) + \\ & \quad \tilde{k}_{21}(\xi_{1,k}, q(\xi_{1,k}, w_k))w_k + O(\varepsilon), \end{cases} \tag{4.101}$$

and a boundary-layer (or quasi steady-state) subsystem as

$$\bar{\xi}_{2,m+1} = \tilde{f}_2(\bar{\xi}_{2,m}, \varepsilon) + \tilde{g}_{21}(\xi_{1,m}, \bar{\xi}_{2,m}, \varepsilon)w_m \tag{4.102}$$

where $m = \lfloor k/\varepsilon \rfloor$ is a stretched-time parameter. This subsystem is guaranteed to be asymptotically stable for $0 < \varepsilon < \varepsilon^\star$ (see Theorem 8.2 in Ref. (Khalil, 1985)) if the original system (4.61) is asymptotically stable.

We can then proceed to redesign the filter (4.65) for the composite system (4.101), (4.102) separately as

$$\widetilde{\mathbf{F}}^{da}_{2c} : \begin{cases} \breve{\xi}_{1,k+1} &= \tilde{f}_1(\breve{\xi}_{1,k}) + \tilde{g}_{11}(\breve{\xi}_{1,k})\breve{w}^\star_{1,k} + \breve{L}_1(\breve{\xi}_{1,k}, y_k)(y_k - \widetilde{h}_{21}(\breve{\xi}_{1,k}) - \widetilde{h}_{22}(\breve{\xi}_{1,k})) \\ \varepsilon\breve{\xi}_{2,k+1} &= \tilde{f}_2(\breve{\xi}_{2,k}, \varepsilon) + \tilde{g}_{21}(\hat{\xi}_k, \varepsilon)\breve{w}^\star_{2,k} + \breve{L}_2(\breve{\xi}_{2,k}, y_k, \varepsilon)(y_k - \tilde{h}_{21}(\breve{\xi}_k, \varepsilon) - \\ & \quad \tilde{h}_{22}(\breve{\xi}_k, \varepsilon)), \end{cases} \tag{4.103}$$

where

$$\widetilde{h}_{21}(\breve{\xi}_{1,k}) = \tilde{h}_{21}(\breve{\xi}_{1,k}, q(\breve{\xi}_{1,k}, \hat{w}^\star_{1,k})), \quad \widetilde{h}_{22}(\breve{\xi}_{1,k}) = \tilde{h}_{21}(\breve{\xi}_{1,k}, q(\breve{\xi}_{1,k}, \hat{w}^\star_{2,k})).$$

Notice also that, ξ_2 cannot be estimated from (4.100) since this is a "quasi-steady-state" approximation. Then, using a similar approximation procedure as in Proposition 4.2.1, we arrive at the following result.

Theorem 4.2.1. *Consider the nonlinear system (4.61) and the $\mathcal{H}_\infty$ estimation problem for this system. Suppose the plant $\mathbf{P_{sp}^{da}}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and zero-input observable. Further, suppose there exists a local diffeomorphism φ that transforms the system to the partially decoupled form (4.64), and Assumption 4.2.1 holds. In addition, suppose for some $\gamma > 0$, there exist C^2 positive-definite functions $\breve{V}_i : \breve{N}_i \times \breve{\Upsilon}_i \to \Re_+$, $i = 1, 2$, locally defined in neighborhoods $\breve{N}_i \times \breve{\Upsilon}_i \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\breve{\xi}_i, y) = (0, 0)$ $i = 1, 2$ respectively, and matrix functions $\breve{L}_i : \breve{N}_i \times \breve{\Upsilon}_i \to \Re^{n_i \times m}$, $\breve{\Upsilon}_i \subset \mathcal{Y}$, $i = 1, 2$ satisfying the pair of DHJIEs:*

$$
\begin{aligned}
&\breve{V}_1(\tilde{f}_1(\breve{\xi}_1), y) + \frac{1}{2\gamma^2}\breve{V}_{1,\breve{\xi}_1}(\tilde{f}_1(\breve{\xi}_1), y)\tilde{g}_{11}(\breve{\xi}_1, q(\breve{\xi}_1, \breve{w}_1^\star))\tilde{g}_{11}^T(\breve{\xi}_1, q(\breve{\xi}_1, \breve{w}_1^\star))\breve{V}_{1,\breve{\xi}_1}^T(\tilde{f}_1(\breve{\xi}_1), y) - \\
&\quad \bar{V}_1(\breve{\xi}_1, y_{k-1}) - \frac{3}{2}(y - \widetilde{h}_{21}(\breve{\xi}_1)) - \widetilde{h}_{22}(\breve{\xi}_1))^T(y - \widetilde{h}_{21}(\breve{\xi}_1) - \widetilde{h}_{22}(\breve{\xi}_1)) = 0, \\
&\quad \breve{V}_1(0, 0) = 0, \qquad\qquad (4.104)\\
&\breve{V}_2(\frac{1}{\varepsilon}\tilde{f}_2(\breve{\xi}_2, \varepsilon), y) + \frac{1}{2\gamma^2}\bar{V}_{2,\breve{\xi}_2}(\frac{1}{\varepsilon}\tilde{f}_2(\breve{\xi}_2, \varepsilon), y)\tilde{g}_{21}(\hat{\xi}, \varepsilon)\tilde{g}_{21}^T(\breve{\xi}, \varepsilon)\breve{V}_{2,\breve{\xi}_2}^T(\frac{1}{\varepsilon}\tilde{f}_2(\breve{\xi}_2, \varepsilon), y) - \\
&\quad \breve{V}_2(\breve{\xi}_2, y_{k-1}) - \frac{3}{2}(y - \tilde{h}_{21}(\breve{\xi}, \varepsilon) - \tilde{h}_{22}(\breve{\xi}, \varepsilon))^T(y - \tilde{h}_{21}(\breve{\xi}, \varepsilon) - h_{22}(\breve{\xi}, \varepsilon)) = 0, \\
&\quad \breve{V}_2(0, 0) = 0 \qquad\qquad (4.105)
\end{aligned}
$$

together with the side-conditions

$$
\begin{aligned}
\breve{w}_1^\star &= \frac{1}{\gamma^2}\tilde{g}_{11}^T(\breve{\xi}_1, q(\xi_1, \breve{w}_1^\star))\breve{V}_{1,\breve{\xi}_1}^T(\tilde{f}_1(\breve{\xi}_1), y) && (4.106)\\
\breve{w}_2^\star &= \frac{1}{\gamma^2}\tilde{g}_{21}^T(\breve{\xi}, \varepsilon)\breve{V}_{2,\hat{\xi}_2}^T(\frac{1}{\varepsilon}\tilde{f}_2(\breve{\xi}_2, \varepsilon), y) && (4.107)\\
\hat{V}_{1,\hat{\xi}_1}(\tilde{f}_1(\breve{\xi}_1))\breve{L}_1^\star(\breve{\xi}_1, y, \varepsilon) &= -(y - \widetilde{h}_{21}(\breve{\xi}_1, \varepsilon) - \widetilde{h}_{22}(\breve{\xi}, \varepsilon))^T && (4.108)\\
\breve{V}_{2,\breve{\xi}_2}^T(\frac{1}{\varepsilon}\tilde{f}_2(\breve{\xi}_2, \varepsilon), y)\breve{L}_2^\star(\breve{\xi}, y, \varepsilon) &= -\varepsilon(y - \tilde{h}_{21}(\breve{\xi}, \varepsilon) - \tilde{h}_{22}(\breve{\xi}))^T && (4.109)
\end{aligned}
$$

Then the filter $\widetilde{\mathbf{F}}_{\mathbf{2c}}^{\mathbf{da}}$ solves the $\mathcal{H}_\infty$ filtering problem for the system locally in $\cup\breve{N}_i$.

Proof: We define separately two Hamiltonian functions $H_i : \mathcal{X} \times \mathcal{W} \times \mathcal{Y} \times \Re^{n_i \times m} \times \Re \to$

$\Re$, $i = 1, 2$ for each of the two separate components of the filter (4.103). Then, the rest of the proof follows along the same lines as Proposition 4.2.1. $\square$

Remark 4.2.2. *Comparing (4.104), (4.108) with (4.84), (4.85), we see that the two reduced-order filter approximations are similar. Moreover, notice that $\breve{\xi}_1$ appearing in (4.109), (4.105) is not considered as an additional variable, because it is assumed to be known from (4.103a),(4.108) respectively, and is therefore regarded as a parameter. In addition, we observe that, the DHJIE (4.104) is implicit in $\breve{w}_1^\star$, and therefore, some sort of approximation is required in order to obtain an explicit solution.*

Remark 4.2.3. *Notice also that, in the determination of $\breve{w}_1^\star$, we assume $\bar{\xi}_2 = q(\xi_1, w)$ is frozen in the Hamiltonian H_2, and therefore the contribution to $\breve{w}_1^\star$ from $\tilde{g}_{11}(.,.)$, $\widetilde{h}_{21}(.,.)$ is neglected.*

We can similarly specialize the result of Theorem 4.2.1 to the discrete-time linear system (4.87) in the following corollary.

Corollary 4.2.2. *Consider the DLSPS (4.87) and the $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\mathbf{P}_{sp}^l$ is locally asymptotically stable about the equilibrium-point $x = 0$ and observable. Suppose further, it is transformable to the form (4.89) and Assumption 4.2.1 is satisfied, i.e., $\tilde{A}_2$ is nonsingular. In addition, suppose for some $\gamma > 0$ there exist symmetric positive-definite matrices $\breve{P}_i \in \Re^{n_i \times n_i}$, $\breve{Q}_i \in \Re^{m \times m}$, and matrix $\breve{L}_i \in \Re^{n_i \times m}$, $i = 1, 2$ satisfying the following LMIs*

$$\begin{bmatrix} \tilde{A}_1^T \breve{P}_1 \tilde{A}_1 - \breve{P}_1 - 3\widetilde{C}_{21}^T \widetilde{C}_{21} & \tilde{A}_1^T \breve{P}_1 \widetilde{B}_{11} & 3\widetilde{C}_{21}^T & 0 \\ \widetilde{B}_{11}^T \breve{P}_1 \tilde{A}_1 & -\gamma^2 I & 0 & 0 \\ 3\widetilde{C}_{21} & 0 & \breve{Q}_1 - 3I & 0 \\ 0 & 0 & 0 & -\breve{Q} \end{bmatrix} \leq 0$$

$$\begin{bmatrix} -3\tilde{C}_{21}^T \tilde{C}_{21} & -3\tilde{C}_{21}^T \tilde{C}_{22} & 0 & 3\tilde{C}_{21}^T & 0 \\ -3\tilde{C}_{22}^T \tilde{C}_{21} & \tilde{A}_2^T \breve{P}_2 \tilde{A}_2 - \breve{P}_2 - 3\tilde{C}_{22}^T \tilde{C}_{22} & \tilde{A}_2^T \breve{P}_2 \tilde{B}_{21} & 3\tilde{C}_{22}^T & 0 \\ 0 & \tilde{B}_{21}^T \breve{P}_2 \tilde{A}_2 & \gamma^2 \varepsilon^2 I & 0 & 0 \\ 3\tilde{C}_{21} & 3\tilde{C}_{22} & 0 & \breve{Q}_2 - 3I - \breve{R}_2 & 0 \\ 0 & 0 & 0 & 0 & -\breve{Q}_2 \end{bmatrix} \leq 0$$

$$\begin{bmatrix} 0 & \frac{1}{2}(\tilde{A}_1^T \breve{P}_1 \breve{L}_1 - \widetilde{C}_{21}^T) \\ \frac{1}{2}(\tilde{A}_1^T \hat{P}_1 \breve{L}_1 - \widetilde{C}_{21}^T)^T & (1-\delta_3)I \end{bmatrix} \le 0$$

$$\begin{bmatrix} 0 & 0 & -\frac{1}{2}\tilde{C}_{21}^T \\ 0 & 0 & \frac{1}{2\varepsilon^2}(\tilde{A}_2^T \breve{P}_2 \breve{L}_2 - \tilde{C}_{22}^T) \\ -\frac{1}{2}\tilde{C}_{21} & \frac{1}{2\varepsilon^2}(\tilde{A}_2^T \breve{P}_2 \breve{L}_2 - \tilde{C}_{22}^T)^T & (1-\delta_4)I \end{bmatrix} \le 0$$

for some numbers $\delta_3, \delta_4 > 0$ *and where*

$$\widetilde{B}_{11} = \tilde{B}_{11} + \tilde{C}_{22}\tilde{A}_2^{-1}\tilde{B}_{21}, \quad \widetilde{C}_{21} = \tilde{C}_{21} - \frac{1}{\gamma^2}\tilde{C}_{22}\tilde{A}_2^{-1}\tilde{B}_{21}\tilde{B}_{11}^T\breve{P}_1\tilde{A}_1.$$

Then, the filter $\mathbf{F}_{2c}^{dl}$ *solves the* $\mathcal{H}_\infty$ *filtering problem for the system.*

Proof: We take similarly,

$$\breve{V}_1(\hat{\xi}_1, y) = \frac{1}{2}(\breve{\xi}_1^T \breve{P}_1 \breve{\xi}_1 + y^T \breve{Q}_1 y)$$

$$\breve{V}_2(\hat{\xi}_2, y) = \frac{1}{2}(\breve{\xi}_2^T \breve{P}_2 \breve{\xi}_2 + y^T \breve{Q}_2 y)$$

and apply the result of the Theorem. □

4.2.3 Discrete-time Aggregate $\mathcal{H}_\infty$ Filters

Similarly, in the absence of the coordinate transformation, φ, discussed in the previous subsection, a filter has to be designed to solve the problem for the aggregate system (4.61). We discuss this class of filters in this subsection. Accordingly, consider the following class of filters:

$$\mathbf{F}_{3ag}^{da} : \left\{ \begin{array}{rcl} \grave{x}_{1,k+1} & = & f_1(\grave{x}_k) + g_{11}(\grave{x}_k)\grave{w}_k^\star + \grave{L}_1(\grave{x}_k, y_k, \varepsilon)[y_k - \\ & & h_{21}(\grave{x}_{1,k}) - h_{22}(\grave{x}_{2,k})]; \quad \grave{x}_1(k_0) = \bar{x}^{10} \\ \varepsilon\grave{x}_{2,k+1} & = & f_2(\grave{x}_k, \varepsilon) + g_{21}(\grave{x}_k)\grave{w}_k^\star + \grave{L}_2(\grave{x}_k, y_k, \varepsilon)[y_k - \\ & & h_{21}(\grave{x}_{1,k}) - h_{22}(\grave{x}_{2,k})]; \quad \grave{x}_2(k_0) = \bar{x}^{20} \\ \grave{z}_k & = & y_k - h_{21}(\grave{x}_{1,k}) - h_{22}(\grave{x}_{2,k}), \end{array} \right. \qquad (4.110)$$

where $\grave{L}_1, \grave{L}_2 \in \Re^{n\times m}$ are the filter gains, and $\grave{z}$ is the new penalty variable. We can repeat the same kind of derivation above to arrive at the following.

Theorem 4.2.2. *Consider the nonlinear system (4.61) and the $\mathcal{H}_\infty$ estimation problem for this system. Suppose the plant $\mathbf{P}^{\mathbf{da}}_{\mathbf{sp}}$ is locally asymptotically stable about the equilibrium-point $x = 0$, and zero-input observable. Further, suppose there exist a C^1 positive-definite function $\grave{V} : \grave{N} \times \grave{\Upsilon} \to \Re_+$, locally defined in a neighborhood $\grave{N} \times \grave{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\grave{x}_1, \grave{x}_2, y) = (0, 0, 0)$, and matrix functions $\grave{L}_i : \grave{N} \times \grave{\Upsilon} \to \Re^{n_i\times m}$, $i = 1, 2$, satisfying the DHJIE:*

$$\begin{aligned}
&\grave{V}(f_1(\grave{x}), \frac{1}{\varepsilon}f_2(\grave{x},\varepsilon), y) - \grave{V}(\grave{x}, y_{k-1}) + \frac{1}{2\gamma^2}[\ \grave{V}_{\grave{x}_1}(f_1(\grave{x}), \tfrac{1}{\varepsilon}f_2(x,\varepsilon), y) \quad \grave{V}_{\grave{x}_2}(f_1(\grave{x}), \tfrac{1}{\varepsilon}f_2(\grave{x},\varepsilon), y)\ ] \times \\
&\qquad \begin{bmatrix} g_{11}(\grave{x})g_{11}^T(\grave{x}) & \frac{1}{\varepsilon}g_{11}(\grave{x})g_{21}^T(\grave{x}) \\ \frac{1}{\varepsilon}g_{21}(\grave{x})g_{11}^T(\grave{x}) & \frac{1}{\varepsilon^2}g_{21}(\grave{x})g_{21}^T(\grave{x}) \end{bmatrix} \begin{bmatrix} \grave{V}^T_{\grave{x}_1}(f_1(\grave{x}), \frac{1}{\varepsilon}f_2(\grave{x},\varepsilon), y) \\ \grave{V}^T_{\grave{x}_2}(f_1(\grave{x}), \frac{1}{\varepsilon}f_2(\grave{x},\varepsilon), y) \end{bmatrix} - \\
&\frac{3}{2}(y - h_{21}(\grave{x}_1) - h_{22}(\grave{x}_2))^T(y - h_{21}(\grave{x}_1) - h_{22}(\grave{x}_2)) = 0, \quad \hat{V}(0,0) = 0, \qquad (4.111)
\end{aligned}$$

together with the side-conditions

$$\begin{aligned}
\grave{V}_{\grave{x}_1}(f_1(\grave{x}), \frac{1}{\varepsilon}f_2(\grave{x},\varepsilon), y)\grave{L}_1^\star(\grave{x}, y) &= -(y - h_{21}(\grave{x}_1) - h_{22}(\grave{x}_2))^T, \qquad (4.112)\\
\grave{V}_{\grave{x}_2}(f_1(\grave{x}), \frac{1}{\varepsilon}f_2(\grave{x},\varepsilon), y)\grave{L}_2^\star(\grave{x}, y) &= -\varepsilon(y - h_{21}(\grave{x}_1) - h_{22}(\grave{x}_2)). \qquad (4.113)
\end{aligned}$$

Then, the filter $\mathbf{F}^a_{3ag}$ solves the $\mathcal{H}_\infty$ filtering problem for the system locally in $\grave{N}$.

Proof: Proof follows along the same lines as Proposition 4.2.1. □

For the DLSPS (4.87), the Chang transformation φ is always available as given by (4.88). Moreover, the result of Theorem 4.2.2 specialized to the DLSPS is horrendous, in the sense that, the resulting inequalities are not linear and too involved. Thus, it is more useful to consider the reduced-order filter which will be introduced shortly as a special case.

Using similar procedure as outlined in the previous section, we can obtain the limiting behavior of the filter $\mathbf{F}^{a}_{3ag}$ as $\varepsilon \downarrow 0$

$$\bar{\mathbf{F}}^{da}_{5ag} \;:\; \begin{cases} \grave{x}_{1,k+1} \;=\; f_1(\grave{x}_k) + g_{11}(\grave{x}_k)\grave{w}^{\star}_{10,k} + \grave{L}_{10}(\grave{x}_k, y_k)(y_k - h_{21}(\grave{x}_{1,k})); \\ \qquad\qquad \grave{x}_1(k_0) = \bar{x}^{10} \\ \grave{x}_{2,k} \;\to\; 0, \end{cases} \tag{4.114}$$

with

$$\grave{w}^{\star}_{10} = \frac{1}{\gamma^2} g^T_{11}(\grave{x})\grave{V}^T_{\grave{x}_1}(f_1(\grave{x})),$$

and the DHJIE (4.111) reduces to the DHJIE

$$\begin{aligned} &\grave{V}(f_1(\grave{x}_1), y) + \frac{1}{2\gamma^2}\bar{V}_{\grave{x}_1}(f_1(\grave{x}_1), y)g_{11}(\grave{x})g^T_{11}(\grave{x})\grave{V}^T_{\grave{x}_1,y}(f_1(\grave{x})) - \grave{V}(\grave{x}_1, y_{k-1}) - \\ &\frac{3}{2}(y - h_{21}(\grave{x}_1))^T(y - h_{21}(\grave{x}_1)) = 0, \quad \grave{V}(0,0) = 0, \end{aligned} \tag{4.115}$$

together with the side-conditions

$$\grave{V}_{\grave{x}_1}(f_1(\grave{x}_1))\grave{L}^{\star}_{10}(\grave{x}, y) \;=\; -(y - h_{21}(\grave{x}_1))^T \tag{4.116}$$

$$\grave{L}_2(\grave{x}, y) \;\to\; 0. \tag{4.117}$$

Similarly, specializing the above result to the DLSPS (4.87), we obtain the following reduced-order filter

$$\mathbf{F}^{dl}_{6agr} : \left\{ \; \grave{x}_{1,k+1} \;=\; A_1\grave{x}_{1,k} + B_{11}\grave{w}^{\star}_{10,k} + \grave{L}^{\star}_{10}(y_k - \tilde{C}_{21}\grave{x}_{1,k}), \right. \tag{4.118}$$

with

$$\grave{w}^{\star}_{10} = \frac{1}{\gamma^2} B^T_{11}\grave{P}_1 A_1 \grave{x}_1$$

and the DHJIE (4.115) reduces to the LMI

$$\begin{bmatrix} A_1^T \grave{P}_{10}\tilde{A}_1 - \grave{P}_{10} - 3C_{21}^T C_{21} & A_1^T \grave{P}_{10} B_{11} & 3C_{21}^T & 0 \\ B_{11}^T \grave{P}_{10} A_1 & -\gamma^2 I & 0 & 0 \\ 3C_{21} & 0 & \grave{Q}_1 - 3I & 0 \\ 0 & 0 & 0 & -\grave{Q} \end{bmatrix} \leq 0 \tag{4.119}$$

$$\begin{bmatrix} 0 & \frac{1}{2}(A_1^T \grave{P}_{10}\grave{L}_{10} - C_{21}^T) \\ \frac{1}{2}(A_1^T \grave{P}_{10}\grave{L}_{10} - C_{21}^T)^T & (1-\delta_5)I \end{bmatrix} \leq 0 \tag{4.120}$$

for some symmetric positive-definite matrices $\grave{P}_{10}$, $\grave{Q}_{10}$, gain matrix $\grave{L}_{10}$ and some number $\delta_5 > 1$.

Remark 4.2.4. *If the nonlinear system (4.61) is in the standard form, i.e., the equivalent of Assumption 4.2.1 is satisfied, and there exists at least one root* $\bar{x}_2 = \sigma(x_1, w)$ *to the equation*

$$0 = f_2(x_1, x_2) + g_{21}(x_1, x_2)w,$$

then reduced-order filters can also be constructed for the system similar to the result of Subection 4.3.3 and Theorem 4.2.1. Such filters would take the following form

$$\mathbf{F}^a_{7agr} : \begin{cases} \check{x}_{1,k+1} &= f_1(\check{x}_{1,k}, \sigma(\check{x}_1, \check{w}^\star_{1,k})) + g_{11}(\check{x}_1, \sigma(\check{x}_1, \check{w}^\star_{1,k}))\check{w}^\star_{1,k} + \\ & \quad \check{L}_1(\check{x}_{1,k}, y_k, \varepsilon)(y_k - h_{21}(\check{x}_{1,k}) - h_{22}(\sigma(\check{x}_1, \check{w}^\star_{1,k})); \quad \check{x}_1(k_0) = \bar{x}_{10} \\ \varepsilon\check{x}_{2,k+1} &= f_2(\check{x}_k, \varepsilon) + g_{21}(\check{x}_1, \check{x}_2)\check{w}^\star_{2,k} + \\ & \quad \check{L}_2(\check{x}_k, y_k, \varepsilon)(y_k - h_{21}(\check{x}_{1,k}) - h_{22}(\check{x}_{2,k})); \quad \check{x}_2(k_0) = \bar{x}_{20} \\ \check{z}_k &= y_k - h_{21}(\check{x}_{1,k}) - h_{22}(\check{x}_{2,k}). \end{cases} \tag{4.121}$$

However, this filter would fall into the class of decomposition filters, rather than aggregate, and because of this, we shall not discuss it further in this subsection.

In the next section, we consider an example.

4.2.4 Examples

Consider the following singularly-perturbed nonlinear system

$$\begin{aligned} x_{1,k+1} &= x_{1,k}^{\frac{1}{3}} + x_{2,k}^{\frac{1}{2}} + w \\ \varepsilon x_{2,k+1} &= -x_{2,k}^{\frac{1}{2}} - x_{2,k}^{\frac{1}{3}} \\ y_k &= x_{1,k} + x_{2,k} + w. \end{aligned}$$

where $w \in \ell_2[0,\infty)$ is a noise process, $\varepsilon \geq 0$. We construct the aggregate filter $\mathbf{F}^a_{3ag}$ presented in the previous section for the above system. It can be checked that the system is locally observable, and with $\gamma = 1$, the function $\breve{V}(\grave{x}) = \frac{1}{2}(\grave{x}_1^2 + \varepsilon \grave{x}_2^2)$, solves the inequality form of the DHJIE (4.111) corresponding to the system. Subsequently, we calculate the gains of the filter as

$$\grave{L}_1(\grave{x}, y) = -\frac{(y - \grave{x}_1 - \grave{x}_2)}{\grave{x}_1^{\frac{1}{3}} + \grave{x}_2^{\frac{1}{2}}}, \quad \grave{L}_2(\grave{x}, y) = \varepsilon \frac{(y - \grave{x}_1 - \grave{x}_2)}{\grave{x}_2^{\frac{1}{2}} + \grave{x}_2^{\frac{1}{3}}}, \tag{4.122}$$

where the gains $\grave{L}_1$, $\grave{L}_2$ are set equal to zero if $|\grave{x}_1^{\frac{1}{3}} + \grave{x}^{\frac{1}{2}}| < \epsilon$ (small), $|\grave{x}_2^{\frac{1}{2}} + \grave{x}_2^{\frac{1}{3}}| < \epsilon$ (small) to avoid the singularity at the origin $\grave{x} = 0$.

4.3 Conclusion

In this chapter, we have presented a solution to the $\mathcal{H}_\infty$ local filtering problem for affine nonlinear singularly-perturbed systems in both continuous-time and discrete-time. Two main types of filters, namely, decomposition and aggregate filters have been presented, and sufficient conditions for the solvability of the problem using each filter are given in terms of HJIEs. Moreover, for the continuous-time problem and the decomposition filters, the solution to the mixed $\mathcal{H}_2/\mathcal{H}_\infty$ filtering problem has also been presented. While for the discrete-time problem, first-order approximate solutions have been derived.

Furthermore, for each type of filter, reduced-order filters have also been derived as limiting cases of the above filters as the singular parameter $\varepsilon \downarrow 0$. The results have also been specialized to linear systems, in which case the conditions reduce to a system of matrix-inequalities(MI) and LMIs which are computationally efficient to solve. Moreover, it has been shown that, based on the local linear approximation of the nonlinear models, it is possible to find the minimum disturbance attenuation levels γ^* and upper bounds ε^* on the singular parameter ε which guarantee the asymptotic stability of the filters. In addition, examples have been presented to illustrate the approach.

Future efforts would concentrate in finding explicit form for the coordinate transformation discussed in Section 3, and developing computationally efficient algorithms for solving the HJIEs.

CHAPTER 5

$\mathcal{H}_2$ FILTERING FOR NONLINEAR DESCRIPTOR SYSTEMS

In this chapter, we discuss the $\mathcal{H}_2$ or Kalman filtering problem for affine nonlinear descriptor systems. The linear problem has been discussed in several references (Dai, 1989), (Nikoukhah, 1999), (Nikoukhah, 1992), (Zhou, 2008), however, to the best of our knowledge, the filtering problem for more general affine nonlinear descriptor systems has not been discussed in any reference. Therefore, in this chapter we propose to discuss this problem for both continuous-time and discrete-time systems. Two classes of filters will be presented; namely, (i) singular type, and (ii) normal type filters. Moreover, $\mathcal{H}_2$ filtering techniques are useful when the system noise and measurement noise are known to be approximately Gaussian distributed, and are superior to $\mathcal{H}_\infty$ techniques in such applications.

In addition, while the extended Kalman-filter has remained the most popular tool used in this area, it still suffers from the problem of local linearization around the previous estimate, and as such, the convergence of the estimates cannot be guaranteed either empirically or theoretically. On the other hand, the result that we present in this chapter employ the full nonlinear system dynamics, and proof of asymptotic convergence can be established. The chapter is organized as follows. In Section 5.1, we present results for the continuous-time problem, while in Section 5.3, we present corresponding results for the discrete-time problem. Finally, in Section 5.4, we present a short conclusion.

5.1 $\mathcal{H}_2$ Filtering for Continuous-time Nonlinear Descriptor Systems

In this section, we discuss the filtering problem for continuous-time systems, while in the next section, we discuss the corresponding results for discrete-time systems. We begin with the problem definition.

5.1.1 Problem Definition and Preliminaries

The general set-up for studying $\mathcal{H}_2$ filtering problems is shown in Fig. 5.1, where $\mathbf{P}$ is the plant, while $\mathbf{F}$ is the filter. The noise signal $w \in \mathcal{S}$ is in general a bounded spectral signal (e.g. a Gaussian white-noise signal), which belongs to the set $\mathcal{S}$ of bounded spectral signals, while $\tilde{z} \in \mathcal{P}$ is a bounded power signal or $\mathcal{L}_2$ signal, which belongs to the space of bounded power signals. Thus, the induced norm from w to $\tilde{z}$ (the penalty variable to be defined later) is the $\mathcal{L}_2$-norm of the interconnected system $\mathbf{F} \circ \mathbf{P}$, i.e.,

$$\|\mathbf{F} \circ \mathbf{P}\|_{\mathcal{L}_2} \triangleq \sup_{0 \neq w \in \mathcal{S}} \frac{\|\tilde{z}\|_{\mathcal{P}}}{\|w\|_{\mathcal{S}}}, \tag{5.1}$$

and is defined as the $\mathcal{H}_2$-norm of the system for stable plant-filter pair $\mathbf{F} \circ \mathbf{P}$, where the operator $\circ$ implies composition of input-output maps. At the outset, we consider the following affine nonlinear causal descriptor model of the plant which is defined on a manifold $\mathcal{X} \subseteq \Re^n$ with zero control input:

$$\mathbf{P}^a_D : \begin{cases} E\dot{x} &= f(x) + g_1(x)w; \quad x(t_0) = x_0 \\ y &= h_2(x) + k_{21}(x)w \end{cases} \tag{5.2}$$

where $x \in \mathcal{X}$ is the semistate vector; $w \in \mathcal{W} \subset \Re^m$ is an unknown disturbance (or noise) signal, which belongs to the set $\mathcal{W}$ of admissible exogenous inputs; $y \in \mathcal{Y} \subset \Re^m$ is the measured output (or observation) of the system, and belongs to $\mathcal{Y}$, the set of admissible measured-outputs.

The functions $f : \mathcal{X} \to T\mathcal{X}$ [1], $g_1 : \mathcal{X} \to \mathcal{M}^{n_1 \times m}(\mathcal{X})$, where $\mathcal{M}^{i \times j}$ is the ring of $i \times j$ smooth matrices over $\mathcal{X}$, $h_2 : \mathcal{X} \to \Re^m$, and $k_{21} : \mathcal{X} \to \mathcal{M}^{m \times m}(\mathcal{X})$ are real C^∞ functions of x, while $E \in \Re^{n \times n}$ is a constant singular matrix. Furthermore, we assume without any loss of generality that the system (5.2) has an isolated equilibrium-point at $x = 0$ such that $f(0) = 0$, $h_2(0) = 0$. We also assume that there exists at least one solution $x(t, t_0, Ex_0, w)$ $\forall t \in \Re$ for the system for all admissible initial conditions Ex_0, for all $w \in \mathcal{W}$. Further, the initial

[1] For a manifold M, TM and $T^\star M$ are the tangent and cotangent bundles of M.

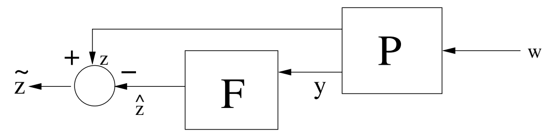

Figure 5.1 Set-up for $\mathcal{H}_2$ Filtering

condition Ex_0 is said to be admissible, if the solution $x(t)$ is unique, impulse-free and smooth for all $[t_0, \infty)$.

In addition, the following standing assumptions will be made on the system.

Assumption 5.1.1. *Let $\bar{x} \in \mathcal{O} \subset \mathcal{X}$, $A = \frac{\partial f}{\partial x}(\bar{x})$. Then, the system (5.2) is admissible implies the following hold:*

1. *the system is locally regular at each $\bar{x} \in \mathcal{O}$ and hence locally solvable, i.e, $det(sE - A) \not\equiv 0$;*
2. *the system is locally impulse-free at each $\bar{x} \in \mathcal{O}$, i.e., $deg(det(sE - A)) = rank(E)$ for all $\bar{x} \in \mathcal{O}$ and al most all $s \in \mathbf{C}$;*
3. *the system is locally asymptotically stable, i.e., (E, A) is Hurwitz at each $\bar{x} \in \mathcal{O}$.*

The standard $\mathcal{H}_2$ local filtering/state-estimation problem is defined as follows.

Definition 5.1.1. *(Standard $\mathcal{H}_2$ Local State Estimation or Filtering Problem). Find a filter, $\mathbf{F}$, for estimating the state $x(t)$ or a function of it, $z = h_1(x)$, from observations $\mathsf{Y}_t \triangleq \{y(\tau) : \tau \leq t\}$, of $y(\tau)$ up to time t, to obtain the estimate*

$$\hat{x}(t) = \mathbf{F}(\mathsf{Y}_t),$$

such that, the $\mathcal{H}_2$-norm from the input w to some suitable penalty function $\tilde{z}$ is locally minimized for all admissible initial conditions $Ex_0 \in \mathcal{O} \subset \mathcal{X}$. Moreover, if the filter solves the problem for all admissible $Ex_0 \in \mathcal{X}$, we say the problem is solved globally.

We shall adopt the following definition of observability (Ozcaldiran, 1992).

Definition 5.1.2. *For the nonlinear system $\mathbf{P}_D^a$, we say that it is locally weakly zero-input observable, if for all states $x_1,\ x_2 \in U \subset \mathcal{X}$ and input $w(.) = 0$, $t > t_0$*

$$y(t; Ex_1(t_0-), w) \equiv y(t; Ex_2(t_0-), w) \Longrightarrow Ex_1(t_0) = Ex_2(t_0); \tag{5.3}$$

the system is said to be locally zero-input observable if

$$y(t; Ex_1(t_0-), w) \equiv y(t; Ex_2(t_0-), w) \Longrightarrow x_1(t_0) = x_2(t_0); \tag{5.4}$$

where $y(., Ex_i(t_0-), w)$, $i = 1, 2$ is the output of the system with the initial condition $Ex_i(t_0-)$; and the system is said to be locally strongly zero-input observable if

$$y(t; Ex_1(t_0-), w) \equiv y(t; Ex_2(t_0-), w) \Longrightarrow x_1(t_0-) = x_2(t_0-). \tag{5.5}$$

Moreover, the system is said to be globally (weakly, strongly) zero-input observable, if it is locally (weakly, strongly) zero-input observable at each $x_0 \in \mathcal{X}$ or $U = \mathcal{X}$.

In the sequel, we shall not distinguish between zero-input observability and strong zero-input observability.

5.1.2 $\mathcal{H}_2$ Singular Filters

In this subsection, we discuss singular filters for the $\mathcal{H}_2$ state estimation problem defined in the previous section. We then discuss normal filters in the next subsection. For this purpose, we assume that the noise signal $w \in \mathcal{W} \subset \mathcal{S}$ is a zero-mean Gaussian white-noise process, i.e.,

$$\mathsf{E}\{w(t)\} = 0, \quad \mathsf{E}\{w(t)w^T(\tau)\} = W\delta(t - \tau).$$

We then consider the following class of singular filters for the system with the optimal noise level set at $\hat{w}^\star = \mathsf{E}w = 0$ in the usual Kalman-Luenberger type structure:

$$\mathbf{F}^a_{DS1} \begin{cases} E\dot{\hat{x}} &= f(\hat{x}) + \hat{L}(\hat{x}, y)(y - h_2(\hat{x})); \quad \hat{x}(t_0) = 0 \\ \tilde{z} &= y - h_2(\hat{x}) \end{cases} \tag{5.6}$$

where $\hat{x} \in \mathcal{X}$ is the filter state, $\hat{L} : \mathcal{X} \times \mathcal{Y} \to \Re^{n\times m}$ is the gain matrix of the filter, and $\tilde{z} \in \Re^m$ is the penalty variable or estimation error.

The problem can then be formulated as a dynamic optimization problem with the following cost functional

$$\begin{aligned} \min_{\hat{L}\in\Re^{n\times m}, w\in\mathcal{S}, x_0=0} \hat{J}(\hat{L}, w) &= \mathsf{E}\left\{\frac{1}{2}\int_{t_0}^{\infty} \{\|\tilde{z}\|^2_W\} dt\right\} = \frac{1}{2}\left\{\|\mathbf{F}^a_S \circ \mathbf{P}^a_D\|^2_{\mathcal{H}_2}\right\}_W \\ & s.t.\ (5.6),\ \text{and with } w = 0, \lim_{t\to\infty}\{\hat{x}(t) - x(t)\} = 0. \end{aligned} \tag{5.7}$$

To solve the above problem, we form the Hamiltonian function $H : T^\star\mathcal{X} \times T^\star\mathcal{Y} \times \mathcal{W} \times \Re^{n\times m} \to \Re$:

$$H(\hat{x}, y, w, \hat{L}, \hat{V}^T_{E\hat{x}}, \hat{V}^T_y) = \hat{V}_{E\hat{x}}(E\hat{x}, y)[f(\hat{x}) + \hat{L}(\hat{x}, y)(y - h_2(\hat{x}))] + \hat{V}_y(E\hat{x}, y)\dot{y} + \tag{5.8}$$

$$\frac{1}{2}\|\tilde{z}\|^2_W \tag{5.9}$$

for some C^1 function $\hat{V} : \mathcal{X} \times \mathcal{Y} \to \Re$, and where $\hat{V}_{E\hat{x}}$ is the row vector of first partial derivatives of $\hat{V}$ with respect to $E\hat{x}$.

Completing the squares now for $\hat{L}$ in the above expression (5.9), we have

$$\begin{aligned} H(\hat{x}, y, w, \hat{L}, \hat{V}^T_{E\hat{x}}, \hat{V}^T_y) &= \hat{V}_{E\hat{x}}(E\hat{x}, y) f(\hat{x}) + \hat{V}_y(E\hat{x}, y)\dot{y} + \\ & \frac{1}{2}\|\hat{L}^T(\hat{x}, y)\hat{V}^T_{E\hat{x}}(E\hat{x}, y) + (y - h_2(\hat{x}))\|^2 - \\ & \frac{1}{2}\hat{V}_{E\hat{x}}(E\hat{x}, y)\hat{L}(\hat{x}, y)\hat{L}^T(\hat{x}, y)\hat{V}^T_{E\hat{x}}(E\hat{x}, y) + \frac{1}{2}\|\tilde{z}\|^2_{(W-I)}. \end{aligned}$$

Thus, setting the optimal gain $\hat{L}^{\star}(\hat{x}, y)$ as

$$\hat{V}_{E\hat{x}}(E\hat{x}, y)\hat{L}^{\star}(\hat{x}, y) = -(y - h_2(\hat{x}))^T, \tag{5.10}$$

minimizes the Hamiltonian (5.9). Finally, setting

$$H(\hat{x}, y, w, \hat{L}, \hat{V}_{E\hat{x}}^T, \hat{V}_y^T) = 0$$

results in the following Hamilton-Jacobi-Bellman equation (HJBE):

$$\begin{aligned}&\hat{V}_{E\hat{x}}(E\hat{x}, y)f(\hat{x}) + \hat{V}_y(E\hat{x}, y)\dot{y} - \frac{1}{2}\hat{V}_{E\hat{x}}(E\hat{x}, y)\hat{L}(\hat{x}, y)\hat{L}^T(\hat{x}, y)\hat{V}_{E\hat{x}}^T(E\hat{x}, y) + \\ &\frac{1}{2}(y - h_2(\hat{x}))^T(W - I)(y - h_2(\hat{x})) = 0, \;\; \hat{V}(0, 0) = 0,\end{aligned} \tag{5.11}$$

or equivalently the HJBE

$$\hat{V}_{E\hat{x}}(E\hat{x}, y)f(\hat{x}) + \hat{V}_y(E\hat{x}, y)\dot{y} + \frac{1}{2}(y - h_2(\hat{x}))^T(W - 2I)(y - h_2(\hat{x})) = 0, \;\; \hat{V}(0, 0) = 0. \tag{5.12}$$

But notice from (5.2), with the measurement noise set at zero,

$$\dot{y} = \tilde{\mathcal{L}}_{f+gw}h_2,$$

where $\tilde{\mathcal{L}}$ is the Lie-derivative operator (Sastry, 1999) in coordinates Ex. Moreover, under certainty-equivalence and with $\hat{w}^{\star} = \mathsf{E}\{w\} = 0$, we have

$$\dot{y} = \tilde{\mathcal{L}}_{f(\hat{x})}h_2(\hat{x}) = \nabla_{E\hat{x}}h_2(\hat{x})f(\hat{x}).$$

Substituting now the above expression in the HJBE (5.12), results in the following formal form of the equation

$$\begin{aligned}&\hat{V}_{E\hat{x}}(E\hat{x}, y)f(\hat{x}) + \hat{V}_y(E\hat{x}, y)\nabla_{E\hat{x}}h_2(\hat{x})f(\hat{x}) + \frac{1}{2}(y - h_2(\hat{x}))^T(W - 2I)(y - h_2(\hat{x})) = 0, \\ &\hat{V}(0, 0) = 0.\end{aligned} \tag{5.13}$$

Consequently, we then have the following result.

Proposition 5.1.1. *Consider the nonlinear system (5.2) and the $\mathcal{H}_2$ filtering problem for this system. Suppose the plant $\mathbf{P_D^a}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and zero-input observable. Further, suppose there exist a C^1 positive-semidefinite function $\hat{V} : \hat{N} \times \hat{\Upsilon} \to \Re_+$ locally defined in a neighborhood $\hat{N} \times \hat{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\hat{\xi}, y) = (0, 0)$, and a matrix function $\hat{L} : \hat{N} \times \hat{\Upsilon} \to \Re^{n\times m}$, satisfying the HJBE (5.13) together with the side-condition (5.10). Then the filter $\mathbf{F}^a_{DS1}$ solves the $\mathcal{H}_2$ filtering problem for the system locally in $\hat{N}$.*

Proof: The optimality of the filter gain $\hat{L}^\star$ has already been shown above. It remains to prove asymptotic convergence of the estimation error vector. Accordingly, let $\hat{V} \geq 0$ be a C^1 solution of the HJBE (5.11) or equivalently (5.12). Then, differentiating this solution along a trajectory of (5.6), with $\hat{L} = \hat{L}^\star$, we get

$$\begin{aligned} \dot{\hat{V}} &= \hat{V}_{E\hat{x}}(E\hat{x}, y)[f(\hat{x}) + \hat{L}^\star(\hat{x}, y)(y - h_2(\hat{x}))] + \hat{V}_y(E\hat{x}, y)\dot{y} \\ &= -\frac{1}{2}\|z\|^2_W, \end{aligned}$$

where the last equality follows from using the HJBE (5.12). Therefore, the filter dynamics is stable, and $V(E\hat{x}, y)$ is non-increasing along a trajectory of (5.6). Further, the condition that $\dot{\hat{V}}(E\hat{x}(t), y(t)) \equiv 0 \ \forall t \geq t_s$ implies that $z \equiv 0$, which further implies that $y = h_2(\hat{x}) \ \forall t \geq t_s$. By the zero-input observability of the system, this implies that $\hat{x} = x \ \forall t \geq t_s$. $\square$

The result of the theorem can be specialized to the linear descriptor system

$$\mathbf{P}^l_D \; : \; \begin{cases} E\dot{x} &= Ax + B_1 w; \quad Ex(t_0) = Ex_0 \\ y &= C_2 x + D_{21} w \end{cases} \tag{5.14}$$

where $E \in \Re^{n\times n}$, $A \in \Re^{n\times n}$, $B_1 \in \Re^{n\times m}$, $C_2 \in \Re^{m\times n}$, $D_{21} \in \Re^{m\times r}$. Assuming without loss of generality that $W = I$, we have the following result.

Corollary 5.1.1. *Consider the linear descriptor system (5.14) and the $\mathcal{H}_2$ filtering problem for this system. Suppose the plant $\mathbf{P}^l_D$ is locally asymptotically stable about the equilibrium-*

point $x = 0$ and observable. Suppose further, there exist symmetric positive-semidefinite matrices $\hat{P} \in \Re^{n\times n}$, $\hat{Q} \in \Re^{m\times m}$, and a matrix $\hat{L} \in \Re^{n\times m}$, satisfying the linear matrix-inequalities

$$\begin{bmatrix} E^T\hat{P}A + A^T\hat{P}E - C_2^TC_2 & C_2^T & 0 \\ C_2 & -I & \hat{Q} \\ 0 & \hat{Q}^T & 0 \end{bmatrix} \leq 0 \tag{5.15}$$

$$\begin{bmatrix} 0 & \frac{1}{2}(E^T\hat{P}\hat{L} - C_2^T) \\ \frac{1}{2}(E^T\hat{P}\hat{L} - C_2^T)^T & (1-\delta_1)I \end{bmatrix} \leq 0 \tag{5.16}$$

for some number $\delta_1 \geq 1$. Then the filter

$$\mathbf{F}^l_{DS1} : \begin{cases} E\dot{\hat{x}} &= A\hat{x} + \hat{L}(y - C_2\hat{x}); \quad E\hat{x}(t_0) = 0 \\ \hat{z} &= C_2\hat{x} \end{cases} \tag{5.17}$$

solves the $\mathcal{H}_2$ estimation problem for the system.

Proof: Take

$$\hat{V} = \frac{1}{2}(\hat{x}^TE^T\hat{P}E\hat{x} + y^T\hat{Q}y), \quad \hat{P} > 0$$

and apply the result of the Proposition. □

Notice however, since the system is inherently constrained, the steady-state error of the estimates may be improved by using a proportional-integral (PI) filter configuration (Gao, 2004), (Koenig, 1995). Thus, we consider the following class of filters:

$$\mathbf{F}^a_{DS2} \begin{cases} E\dot{\check{x}} &= f(\check{x}) + \check{L}_1(\check{x},\xi,y)(y - h_2(\check{x})) + \check{L}_2(\check{x},\xi,y)\xi \\ \dot{\xi} &= y - h_2(\check{x}) \\ \check{z} &= y - h_2(\check{x}) \end{cases} \tag{5.18}$$

where $\check{x} \in \mathcal{X}$ is the filter state, $\xi \in \Re^m \times \Re$ is the integrator state, and $\check{L}_1, \check{L}_2 : \mathcal{X}\times\mathcal{Y} \to \Re^{n\times m}$ are the proportional and integral gain matrices of the filter respectively. Similarly, using manipulations as in Proposition 5.1.1, we can arrive at the following result.

Theorem 5.1.1. *Consider the nonlinear system (5.2) and the $\mathcal{H}_2$ filtering problem for this system. Suppose the plant $\mathbf{P}^{\mathbf{a}}_{\mathbf{D}}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and zero-input observable. Further, suppose there exists a C^1 positive-semidefinite function $\check{V} : \check{N} \times \check{\Xi} \times \check{\Upsilon} \to \Re_+$ locally defined in a neighborhood $\check{N} \times \check{\Xi} \times \hat{\Upsilon} \subset \mathcal{X} \times \mathcal{Y} \times \Re \times \mathcal{Y}$ of the origin $(\check{x}, \xi, y) = (0, 0, 0)$, and matrix functions $\check{L}_1, \check{L}_2 : \check{N} \times \check{\Xi} \times \check{\Upsilon} \to \Re^{n\times m}$, satisfying the HJBE*

$$\check{V}_{E\check{x}}(E\check{x}, \xi, y)f(\check{x}) + \check{V}_y(E\check{x}, \xi, y)\nabla_{E\check{x}}h_2(\check{x})f(\check{x}) + \check{V}_\xi(E\check{x}, \xi, y)(y - h_2(\check{x})) - \xi^T\xi + \\ \tfrac{1}{2}(y - h_2(\check{x}))^T(W - 2I)(y - h_2(\check{x})) = 0, \quad \check{V}(0, 0, 0) = 0. \tag{5.19}$$

together with the side-conditions

$$\check{V}_{E\check{x}}(E\check{x}, \xi, y)\check{L}_1(\check{x}, \xi, y) = -(y - h_2(\check{x}))^T \tag{5.20}$$

$$\check{V}_{E\check{x}}(E\check{x}, \xi, y)\check{L}_2(\check{x}, \xi, y) = -\xi^T. \tag{5.21}$$

Then the filter $\mathbf{F}^a_{DS2}$ solves the $\mathcal{H}_2$ filtering problem for the system locally in $\check{N}$.

In the next section, we consider the design of normal filters for the system.

5.1.3 $\mathcal{H}_2$ Normal Filters

In this subsection, we discuss normal filters for the system (5.2). We shall consider the design of both full-order and reduced-order filters. We start with the full-order filter first, and in this regard, without any loss of generality, we can assume that E is of the form

$$E = \begin{pmatrix} I_{q\times q} & 0 \\ 0 & 0 \end{pmatrix}.$$

This follows from matrix theory and can easily be proven using the singular-value decomposition (SVD) of E. It follows that, the system can be represented in the canonical form of a differential-algebraic system

$$
\bar{\mathbf{P}}_D^a : \begin{cases} \dot{x}_1 = f_1(x) + g_{11}(x)w; \quad x(t_0) = x_0 \\ 0 = f_2(x) + g_{21}(x)w \\ y = h_2(x) + k_{21}(x)w \end{cases} \tag{5.22}
$$

where $dim(x_1) = q$, $f_1(0) = 0$, $f_2(0) = 0$. Then, if we define

$$
\dot{x}_2 = f_2(x) + g_{21}(x)w,
$$

where $\dot{x}_2$ is a fictitious state vector, and $dim(x_2) = n - q$, the system (5.22) can be represented by a normal state-space system as

$$
\widetilde{\mathbf{P}}_D^a : \begin{cases} \dot{x}_1 = f_1(x) + g_{11}(x)w; \quad x_1(t_0) = x_{10} \\ \dot{x}_2 = f_2(x) + g_{21}(x)w; \quad x_2(t_0) = x_{20} \\ y = h_2(x) + k_{21}(x)w. \end{cases} \tag{5.23}
$$

Now define the set

$$
\Omega_o = \{(x_1, x_2) \in \mathcal{X} \mid \dot{x}_2 \equiv 0\}. \tag{5.24}
$$

Then, we have the following system equivalence

$$
\widetilde{\mathbf{P}}_D^a|_{\Omega_o} = \bar{\mathbf{P}}_D^a. \tag{5.25}
$$

Therefore, to estimate the states of the system (5.22), we need to stabilize the system (5.23) about Ω_o and then design a filter for the resulting system. For this purpose, we consider the following class of filters with $\mathsf{E}\{w\} = 0$:

$$
\mathbf{F}_{DN3}^a \begin{cases} \dot{\grave{x}}_1 = f_1(\grave{x}) + \grave{L}_1(\grave{x}, y)(y - h_2(\grave{x})) \\ \dot{\grave{x}}_2 = f_2(\grave{x}) + \tilde{g}_{22}(x)\alpha_2(\grave{x}) + \grave{L}_2(\grave{x}, y)(y - h_2(\grave{x})) \\ \dot{\grave{z}} = y - h_2(\grave{x}), \end{cases} \tag{5.26}
$$

where $\grave{x} \in \mathcal{X}$ is the filter state, $\grave{L}_1 : \mathcal{X} \times \mathcal{Y} \to \Re^{q\times m}$, $\grave{L}_2 : \mathcal{X} \times \mathcal{Y} \to \Re^{n-q\times m}$ are the filter gain matrices, and $g_{22} : \mathcal{X} \to \mathcal{M}^{(n-q)\times p}$ is a gain matrix for the artificial control input $u = \alpha_2(x) \in \Re^p$ required to stabilize the dynamics $\dot{x}_2$ about Ω_o. Accordingly, we make the following assumption.

Assumption 5.1.2. *The pair $\{f_2, g_{22}\}$ is stabilizable, i.e., $\exists$ a control-Lyapunov-function (CLF), $\bar{V} > 0$, such that $\bar{V}_{x_2}(x)(f_2(x) - g_{22}(x)g_{22}^T(x)\bar{V}_{x_2}^T(x)) < 0$.*

Thus, if Assumption 5.1.2 holds, then we can set $\alpha_2(\grave{x}) = -\frac{1}{\varepsilon}g_{22}^T(\grave{x})\bar{V}_{\grave{x}_2}^T(\grave{x})$, where $\varepsilon > 0$ is small, a high-gain feedback (Young, 1977) to constrain the dynamics on Ω_o as fast as possible. Then, we proceed to design the gain matrices $\grave{L}_1$, $\grave{L}_2$ to estimate the states. Consequently, we have the following result.

Proposition 5.1.2. *Consider the nonlinear system (5.22) and the $\mathcal{H}_2$ estimation problem for this system. Suppose the plant $\bar{\mathbf{P}}_{\mathbf{D}}^{\mathbf{a}}$ is locally asymptotically stable about the equilibrium-point $x = 0$, is zero-input observable and satisfies Assumption 5.1.2. Further, suppose there exists a C^1 positive-semidefinite function $\grave{V} : \grave{N} \times \grave{\Upsilon} \to \Re_+$, locally defined in a neighborhood $\grave{N} \times \grave{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\grave{x}_1, \grave{x}_2, y) = (0, 0, 0)$, and matrix functions $\grave{L}_1 : \grave{N} \times \grave{\Upsilon} \to \Re^{q\times m}$, $\grave{L}_2 : \grave{N} \times \grave{\Upsilon} \to \Re^{n-q\times m}$, satisfying the HJBE:*

$$\grave{V}_{\grave{x}_1}(\grave{x}, y)f_1(\grave{x}) + \grave{V}_{\grave{x}_2}(\grave{x}, y)f_2(\grave{x}) + \grave{V}_{\grave{x}_2}(\grave{x}, y)g_{22}(\grave{x})\alpha_2(\grave{x}) + \grave{V}_y(\grave{x}, y)\nabla_{E\grave{x}}h_2(\grave{x})f(\grave{x}) + \frac{1}{2}(y - h_2(\grave{x}))^T(W - 4I)(y - h_2(\grave{x})) = 0, \;\; \grave{V}(0, 0) = 0 \tag{5.27}$$

together with the side-conditions

$$\grave{V}_{\grave{x}_1}(\grave{x}, y)\grave{L}_1(\grave{x}, y) = -(y - h_2(\grave{x}))^T \tag{5.28}$$

$$\grave{V}_{\grave{x}_2}(\grave{x}, y)\grave{L}_2(\grave{x}, y) = -(y - h_2(\grave{x}))^T. \tag{5.29}$$

Then, the filter $\mathbf{F}_{DN3}^a$ solves the $\mathcal{H}_2$-filtering problem for the system locally in $\grave{N}$.

Proof: Follows along same lines as Proposition 5.1.1.

Remark 5.1.1. *Notice the addition of the high-gain feedback $u = \alpha_2(\grave{x})$, transforms the filter $\mathbf{F}^a_{DN3}$ to a singularly-perturbed system (Young, 1977) with a slow subsystem governed by the dynamics $\dot{x}_1$, and a fast subsystem governed by the x_2-dynamics. This design philosophy is not a coincidence, since descriptor systems are intimately related to singularly-perturbed system. This also suggests an alternative approach to the filter design problem, by considering a singularly-perturbed model of the system (5.22) as*

$$\tilde{\mathbf{P}}^a_{\varepsilon D} : \begin{cases} \dot{x}_1 &= f_1(x) + g_{11}(x)w; \;\; x(t_0) = x_0 \\ \varepsilon\dot{x}_2 &= f_2(x) + g_{21}(x)w, \\ y &= h_2(x) + k_{21}(x)w, \end{cases} \tag{5.30}$$

where $\varepsilon > 0$ is a small parameter, and designing a normal filter for this equivalent system (Aliyu, 2011a). Notice in this case, as $\varepsilon \downarrow 0$, the model (5.30) reduces to the original model (5.22).

Remark 5.1.2. *A common HJBE-CLF can also be utilized in the above design procedure. This can be achieved optimally if we take*

$$\begin{aligned} \alpha_2(\grave{x}) &= -\frac{1}{\varepsilon} g_{22}^T(\grave{x}) \bar{V}^T_{\grave{x}_2}(\grave{x}, y), \\ \bar{V}_{\grave{x}_1}(\grave{x}, y)\grave{L}_1(\grave{x}, y) &= -(y - h_2(\grave{x}))^T, \\ \bar{V}_{\grave{x}_2}(\grave{x}, y)\grave{L}_2(\grave{x}, y) &= -(y - h_2(\grave{x}))^T, \end{aligned}$$

where $\bar{V}$ is a C^1 solution of the following HJBE

$$\bar{V}_{\grave{x}_1}(\grave{x}, y) f_1(\grave{x}) + \bar{V}_{\grave{x}_2}(\grave{x}, y) f_2(\grave{x}) - \tfrac{1}{\varepsilon}\bar{V}_{\grave{x}_2}(\grave{x}, y) g_{22}(\grave{x}) g_{22}^T(\grave{x}) \bar{V}^T_{\grave{x}_2}(\grave{x}, y) +$$
$$\grave{V}_y(\grave{x}, y)\nabla_{E\grave{x}} h_2(\grave{x}) f(\grave{x}) + \tfrac{1}{2}(y - h_2(\grave{x}))^T (W - 4I)(y - h_2(\grave{x})) = 0, \quad \bar{V}(0,0) = 0.$$

Next, we consider a reduced-order normal filter design. Accordingly, partition the state-vector x comformably with $rank(E) = q$ as $x = (x_1^T \;\; x_2^T)^T$ with $dim(x_1) = q$, $dim(x_2) = n - q$

and the state equations as

$$\breve{\mathbf{P}}^a_D : \begin{cases} \dot{x}_1 &= f_1(x_1,x_2) + g_{11}(x_1,x_2)w; \;\; x_1(t_0) = x_{10} \\ 0 &= f_2(x_1,x_2) + g_{21}(x_1,x_2)w; \;\; x_2(t_0) = x_{20} \\ y &= h_2(x) + k_{21}(x)w. \end{cases} \tag{5.31}$$

Then we make the following assumption.

Assumption 5.1.3. *The system is in the standard-form, i.e., the Jacobian matrix* $f_{2,x_2}(x_1,x_2)$ *is nonsigular in an open neighborhood* $\widetilde{U}$ *of* $(0,0)$ *and* $g_{21}(0,0) \neq 0$.

If Assumption 5.1.3 holds, then by the Implicit-function Theorem (Sastry, 1999), there exists a unique C^1 function $\phi : \Re^q \times \mathcal{W} \to \Re^{n-q}$ and a solution

$$\bar{x}_2 = \phi(x_1, w)$$

to equation (5.31b). Thus, the system can be locally represented in $\widetilde{U}$ as the reduced-order system

$$\bar{\mathbf{P}}^a_{rD} : \begin{cases} \dot{x}_1 &= f_1(x_1,\phi(x_1,w)) + g_{11}(x_1,\phi(x_1,w))w; \;\; x_1(t_0) = x_{10} \\ y &= h_2(x_1,\phi(x_1,w)) + k_{21}(x_1,\phi(x_1,w))w. \end{cases} \tag{5.32}$$

We can then design a normal filter of the form

$$\mathbf{F}^a_{DrN4} \begin{cases} \dot{\breve{x}}_1 &= f_1(\breve{x}_1,\phi(\breve{x}_1,0)) + \breve{L}(\breve{x}_1,\phi(\breve{x}_1,0),y)(y - h_2(\breve{x}_1,\phi(\breve{x}_1,0))); \\ & \breve{x}_1(t_0) = \mathsf{E}\{x_{10}\} \\ \breve{z} &= y - h_2(\breve{x}_1,\phi(\breve{x}_1,0)) \end{cases} \tag{5.33}$$

for the system, and consequently, we have the following result.

Theorem 5.1.2. *Consider the nonlinear system (5.22) and the* $\mathcal{H}_2$ *filtering problem for this system. Suppose for the plant* $\bar{\mathbf{P}}^{\mathbf{a}}_{\mathbf{D}}$ *is locally asymptotically stable about the equilibrium-point* $x = 0$, *zero-input observable and Assumption 5.1.3 holds for the system. Further, suppose there exists a* C^1 *positive-semidefinite function* $\breve{V} : \breve{N} \times \breve{\Upsilon} \to \Re_+$, *locally defined* *in a neighborhood* $\breve{N} \times \breve{\Upsilon} \subset \widetilde{U} \times \mathcal{Y}$ *of the origin* $(\breve{x}_1, y) = (0,0)$*, and a matrix function* $\breve{L} : \breve{N} \times \breve{\Upsilon} \to \Re^{q\times m}$*, satisfying the HJBE:*

$$\begin{aligned} &\breve{V}_{\breve{x}_1}(\breve{x}_1, y) f_1(\breve{x}_1, \phi(\breve{x}_1, 0)) + \breve{V}_y(\breve{x}_1, y) \nabla_{E\breve{x}} h_2(\breve{x}_1, \phi(\breve{x}_1, 0)) f_1(\breve{x}_1, \phi(\breve{x}_1, 0)) + \\ &\tfrac{1}{2}(y - h_2(\breve{x}_1, \phi(\breve{x}_1, 0)))^T (W - 2I)(y - h_2(\breve{x}_1, \phi(\breve{x}_1, 0))) = 0, \ \breve{V}(0,0) = 0, \end{aligned} \tag{5.34}$$

together with the side-condition

$$\breve{V}_{\breve{x}_1}(\breve{x}, y)\breve{L}(\breve{x}_1, y) = -(y - h_2(\breve{x}_1, \phi(\breve{x}_1, 0)))^T. \tag{5.35}$$

Then, the filter $\mathbf{F}^a_{DrN4}$ *solves the* $\mathcal{H}_2$ *filtering problem for the system locally in* $\breve{N}$*.*

Proof: Follows along same lines as Proposition 5.1.1.

Similarly, we can specialize the result of Theorem 5.1.2 to the linear system (5.14). The system can be rewritten in the form (5.22) as

$$\mathbf{P}^l_D \ : \ \left\{ \begin{array}{rcl} \dot{x} &=& A_1 x_1 + A_{12} x_2 + B_{11} w; \ \ x_1(t_0) = x_{10} \\ 0 &=& A_{21} x_1 + A_2 x_2 + B_{21} w; \ \ \ x_2(t_0) = x_{20} \\ y &=& C_{21} x_1 + C_{22} x_2 + D_{21} w. \end{array} \right. \tag{5.36}$$

Then, if A_2 is nonsingular (Assumption 5.1.3) we can solve for x_2 in equation (5.36(b)) to get

$$\bar{x}_2 = -A_2^{-1}(A_{21} x_1 + B_{21} w),$$

and the filter (5.33) takes the following form

$$\mathbf{F}^l_{DrN4} \left\{ \begin{array}{rcl} \dot{\breve{x}}_1 &=& (A_1 - A_2^{-1} A_{21})\breve{x}_1 + \breve{L}[y - (C_{21} - C_{22} A_2^{-1} A_{21})\breve{x}_1]; \\ && \breve{x}_1(t_0) = \mathsf{E}\{x_{10}\} \\ \breve{z} &=& y - (C_{21} - C_{22} A_2^{-1} A_{21})\breve{x}_1. \end{array} \right. \tag{5.37}$$

Then, we have the following corollary if we again assume that $W = I$ without loss of generality.

Corollary 5.1.2. *Consider the linear descriptor system (5.14) and the $\mathcal{H}_2$-filtering problem for this system. Suppose the plant $\mathbf{P}_D^l$ is locally asymptotically stable about the equilibrium-point $x=0$, Assumption 5.1.3 holds and the plant is zero-input observable. Suppose further, there exist symmetric positive-semidefinite matrices $\breve{P} \in \Re^{q\times q}$, $\breve{Q} \in \Re^{m\times m}$, and a matrix $\breve{L} \in \Re^{n\times m}$, satisfying the LMIs:*

$$\begin{bmatrix} \tilde{A}_1^T\breve{P} + \breve{P}\tilde{A}_1 - \tilde{C}_2^T\tilde{C}_2^T & \tilde{C}_2^T & 0 \\ \tilde{C}_2 & -I & \breve{Q} \\ 0 & \breve{Q} & 0 \end{bmatrix} \leq 0 \tag{5.38}$$

$$\begin{bmatrix} 0 & \frac{1}{2}(\breve{P}\breve{L} - \tilde{C}_2^T) \\ \frac{1}{2}(\breve{P}\breve{L} - \tilde{C}_2^T)^T & (1-\delta_3)I \end{bmatrix} \leq 0 \tag{5.39}$$

for some $\delta_3 \geq 1$, where $\tilde{A}_1 = (A_1 - A_2^{-1}A_{21})$, $\tilde{C}_2 = (C_{21} - C_{22}A_2^{-1}A_{21})$. Then, the filter (5.37) solves the $\mathcal{H}_2$-filtering problem for the system.

Proof Take

$$\breve{V}(\breve{x}) = \frac{1}{2}(\breve{x}_1^T\breve{P}\breve{x}_1 + y^T\breve{Q}y)$$

and apply the result of the Theorem. □

5.1.4 The General case

In this subsection, we consider the filtering problem for the more general class of affine descriptor systems in which $E = E(x) \in \mathcal{M}^{n\times n}(\mathcal{X})$ is a matrix function of x, and can be represented as

$$\mathbf{P}_{DG}^a : \begin{cases} E(x)\dot{x} &= f(x) + g_1(x)w; \quad x(t_0) = x_0 \\ y &= h_2(x) + k_{21}(x)w, \end{cases} \tag{5.40}$$

where *minimum rank*$(E(x)) = q$ for all $x \in \mathcal{X}$, $E(0) = 0$, and all the other variables and functions have their previous meanings and dimensions. We also have the following modified definition of regularity for the system (Zimmer, 1997).

Definition 5.1.3. *The system (5.40) is regular if and only if, there exists an embedded submanifold (Boothby, 1975) $N \subset \mathcal{X}$ and a vector-field $f^{\#}$ such that, every solution of $\dot{x} = f^{\#}(x)$, $x \in N$, is also a solution of (5.40), and vice-versa.*

We first consider the design of a singular filter for the above system. Accordingly, consider a filter of the form (5.6) for the system defined as

$$\mathbf{F}^{a}_{DS5} \begin{cases} E(x^{\flat})\dot{x}^{\flat} &= f(x^{\flat}) + L^{\flat}(x^{\flat}, y)(y - h_2(x^{\flat})) \\ z^{\flat} &= y - h_2(x^{\flat}), \end{cases} \tag{5.41}$$

where $L^{\flat} \in \Re^{n\times m}$ is the gain of the filter. Suppose also the following assumption holds.

Assumption 5.1.4. *There exists a vector-field $e(x) = (e_1(x), \ldots, e_n(x))^T$ such that*

$$E(x) = \frac{\partial e}{\partial x}(x), \quad e(0) = 0.$$

Remark 5.1.3. *Notice that, $e(x)$ cannot in general be obtained by line-integration of the rows of $E(x)$.*

Then we have the following result.

Theorem 5.1.3. *Consider the nonlinear system (5.32) and the $\mathcal{H}_2$ state estimation problem for this system. Suppose for the plant ${\mathbf{P^a}}_{DG}$ is locally asymptotically stable about the equilibrium-point $x = 0$, and zero-input observable. Further, suppose Assumption 2.4.1 holds, there exists a C^1 positive-semidefinite function $V^{\flat} : TN^{\flat} \times \Upsilon^{\flat} \to \Re_+$, locally defined in a neighborhood $TN^{\flat} \times \Upsilon^{\flat} \subset T\mathcal{X} \times \mathcal{Y}$ of the origin $(e(x), y) = (0, 0)$, and a matrix function $L^{\flat} : N^{\flat} \times \Upsilon^{\flat} \to \Re^{n\times m}$, satisfying the HJBE:*

$$\begin{aligned} &V^{\flat}_{e(x^{\flat})}(e(x^{\flat}), y)f(x^{\flat}) + V^{\flat}_{y}(e(x^{\flat}), y)\nabla_{e(x^{\flat})}h_2(x^{\flat})f(x^{\flat}) + \\ &\frac{1}{2}(y - h_2(x^{\flat}))^T(W - 2I)(y - h_2(x^{\flat})) = 0, \quad V^{\flat}(0, 0) = 0, \end{aligned} \tag{5.42}$$

together with the side-conditions

$$V^{\flat}_{e(x^{\flat})}(e(x^{\flat}),y)L^{\flat}(x^{\flat},y) = -(y-h_2(x^{\flat}))^T. \tag{5.43}$$

Then, the filter $\mathbf{F}^a_{DS5}$ *solves the* $\mathcal{H}_2$ *local filtering problem for the system in* $N^{\flat}$.

Proof: Let $V^{\flat} \geq 0$ be a solution of the HJBE (5.34), and consider the time-derivative of this function along a trajectory of (5.41)

$$\begin{aligned} \dot{V}^{\flat} &= V^{\flat}_{e(x^{\flat})}(e(x^{\flat}),y)E(x^{\flat})\dot{x}^{\flat} + V^{\flat}_y(e(x^{\flat}),y)\dot{y} \\ &= V^{\flat}_{e(x^{\flat})}(e(x^{\flat}),y)[f(x^{\flat}) + L^{\flat}(x^{\flat},y)(y-h_2(x^{\flat}))] + V^{\flat}_y(e(x^{\flat}),y)\dot{y} \\ &= -\frac{1}{2}\|z^{\flat}\|^2_W \end{aligned}$$

where the last equality follows from using the HJBE (5.34). The rest of the proof then follows along the same lines as Proposition 5.1.1. $\square$

A normal filter for the system can also be designed. If $rank(E(x)) = q$ is constant for all $x \in \widetilde{\Upsilon} \subset \mathcal{X}$. Then, it can be shown (Zimmer, 1997) that, there exists a nonsingular transformation $T : \widetilde{\Upsilon} \to \mathcal{M}^{n\times n}(\mathcal{X})$ such that

$$T(x)E(x) = \begin{pmatrix} E_1(x) \\ 0 \end{pmatrix}, \quad T(x)f(x) = \begin{pmatrix} \tilde{f}_1(x) \\ \tilde{f}_2(x) \end{pmatrix},$$

where $E_1 \in \mathcal{M}^{q\times q}(\widetilde{\Upsilon})$ is nonsingular on $\widetilde{\Upsilon}$, and the system (5.40) can similarly be represented in this coordinates as

$$\widetilde{\mathbf{P}}^{\mathbf{a}}_{\mathbf{DG}} : \left\{ \begin{aligned} \dot{x}_1 &= E_1^{-1}(x)\tilde{f}_1(x_1,x_2) + E_1^{-1}(x)\tilde{g}_{11}(x_1,x_2)w; \quad x_1(t_0) = x_0 \\ 0 &= \tilde{f}_2(x_1,x_2) + \tilde{g}_{21}(x_1,x_2)w; \quad x_2(t_0) = x_{20} \\ y &= h_2(x) + k_{21}(x)w, \end{aligned} \right. \tag{5.44}$$

where $\begin{pmatrix} \tilde{g}_{11}(x) \\ \tilde{g}_{21}(x) \end{pmatrix} = T(x)g_1(x)$. Then, a normal filter can be designed for the above system using the procedure outlined in Subsection 5.1.3 and Proposition 5.1.2. Similarly, a reduced-order filter for the system can also be designed as in Theorem 5.1.2 if the equivalent of Assumption 5.1.3 is satisfied for the system. This would also circumvent the problem of satisfying Assumption 5.1.4.

In the next section, we consider an example.

5.1.5 Examples

Consider the following simple nonlinear differential-algebraic system

$$\dot{x}_1 = -x_1^3 + x_2 + x_1^2 w_0 \tag{5.45}$$

$$0 = -x_1 - x_2 + sin(x_1) w_0 \tag{5.46}$$

$$y = x_1 + x_2 + w, \tag{5.47}$$

where w_0 is a uniformly distributed noise process. We find the gain for the singular filter $\mathbf{F}^a_{DS1}$ presented in Subsection 2.2. It can be checked that the system is locally observable, and the function $\hat{V}(\hat{x}) = \frac{1}{2}\hat{x}_1^2$, solves the inequality form of the HJBE (5.12) for the system. Subsequently, we calculate the gain of the filter as

$$\hat{L}(\grave{x}_1, y) = -\frac{(y - \hat{x}_1 - \hat{x}_2)}{\hat{x}_1},$$

where $\hat{L}$ is set equal zero if $|\hat{x}_1| < \epsilon$ (ϵ small) to avoid the singularity at $\hat{x} = 0$.

Figures 5.2 and 5.3 show the result of the simulation with the above filter. In Figure 5.2, the noise variance was set to 0.2 while in Figure 5.3 it was set to 0.5. The result of the simulations show good convergence with unknown system initial condition.

Similarly, we can determine the reduced-order filter gain (5.33) for the above system. Notice that the system also satisfies Assumption 5.1.3, thus we can solve equation (5.47) for x_2 to

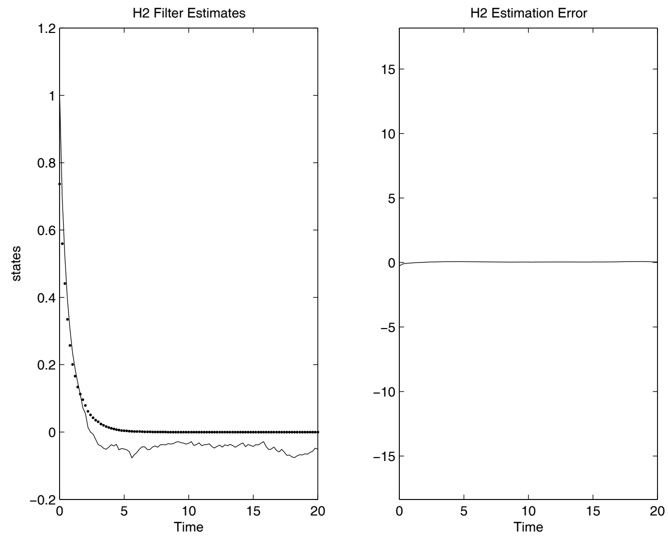

Figure 5.2 $\mathcal{H}_2$ singular filter performance with unknown initial condition and noise variance 0.2

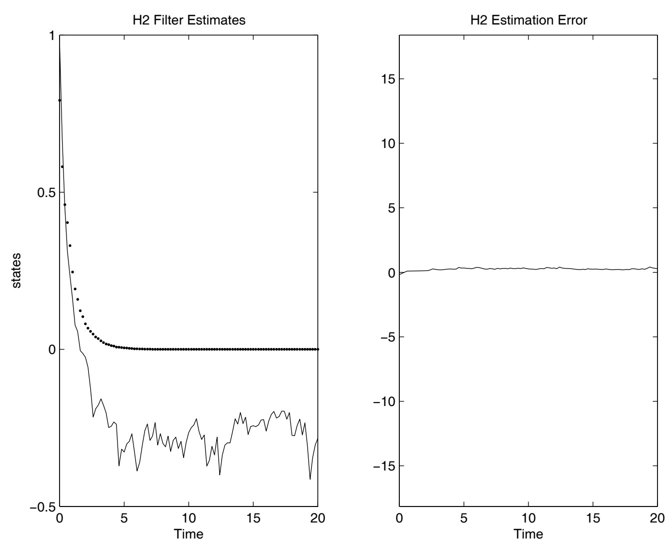

Figure 5.3 $\mathcal{H}_2$ singular filter performance with unknown initial condition and noise variance 0.5

get $\bar{x}_2 = -x_1 + \sin(x_1)w_0$, and substituting in (5.45), we get the reduced system

$$\dot{x}_1 = -x_1^3 - x_1 - sin(x_1)w_0 + x_1^2 w_0$$

which is locally asymptotically stable about $x = 0$. Then, it can be checked that, the function $\breve{V}(x) = \frac{1}{2}\breve{x}_1^2$ solves the HJBE (5.34), and consequently, we have the filter gain

$$\breve{L}_1(\dot{x}_1, y) = -\frac{y}{\breve{x}_1},$$

where again $\breve{L}(\breve{x}_1, y)$ is set equal to zero if $|x_1| < \epsilon$ small. The result of the simulation with this normal or reduced-order filter is very much the same as the singular filter shown above, and hence it is omitted.

5.2 $\mathcal{H}_2$ Filtering for Discrete-time Descriptor Systems

In this section, we present the discrete-time counterpart $\mathcal{H}_2$ filtering results presented in the previous section for affine nonlinear descriptor systems. We shall similarly present two classes of filters, namely, (i) singular; and (ii) normal filters.

5.2.1 Problem Definition and Preliminaries

Again, the set-up for this case is shown in Fig. 5.4, where $\mathbf{P}_k$ is the plant, while $\mathbf{F}_k$ is the filter. The noise signal $w \in \mathcal{S}'$ is in general a bounded spectral signal (e.g. a Gaussian white-noise signal) which belongs to the set $\mathcal{S}'$ of bounded spectral signals, while $\tilde{z} \in \mathcal{P}'$, is a bounded power signal or ℓ_2 signal. Thus, the induced norm from w to $\tilde{z}$ (the penalty variable to be defined later) is the ℓ_2-norm of the interconnected system $\mathbf{F}_k \circ \mathbf{P}_k$, where the operator $\circ$ implies composition of input-output maps, i.e.,

$$\|\mathbf{F}_k \circ \mathbf{P}_k\|_{\ell_2} \triangleq \sup_{0 \neq w \in \mathcal{S}'} \frac{\|\tilde{z}\|_{\mathcal{P}'}}{\|w\|_{\mathcal{S}'}}, \tag{5.48}$$

$$\mathcal{P}' \triangleq \{w:\ w \in \ell_\infty,\ R_{ww}(k),\ S_{ww}(j\omega) \text{ exist for all } k \text{ and all } \omega \text{ resp.},\ \|w\|_{\mathcal{P}'} < \infty\},$$
$$\mathcal{S}' \triangleq \{w:\ w \in \ell_\infty,\ R_{ww}(k),\ S_{ww}(j\omega) \text{ exist for all } k \text{ and all } \omega \text{ resp.},\ \|S_{ww}(j\omega)\|_\infty < \infty\},$$

$$\|z\|^2_{\mathcal{P}'} \triangleq \lim_{K\to\infty} \frac{1}{2K} \sum_{k=-K}^{K} \|z_k\|^2,$$

$$\|w\|_{\mathcal{S}'} = \sqrt{\|S_{ww}(j\omega)\|_\infty} = \sqrt{\sup_w \|S_{ww}(j\omega)\|},$$

and R_{ww}, $S_{ww}(j\omega)$ are the autocorrelation and power spectral density matrices of w. Notice also that, $\|(.)\|_{\mathcal{P}'}$ is a seminorm. In addition, if the plant is stable, we replace the induced ℓ_∞-norm above by the equivalent $\mathcal{H}_\infty$ subspace norms.

At the outset, we consider the following affine nonlinear causal descriptor model of the plant which is defined on $\mathcal{X} \subseteq \Re^n$ with zero control input:

$$\mathbf{P}^{ad}_D : \begin{cases} Ex_{k+1} = f(x_k) + g_1(x_k)w_k; \quad x(k_0) = x^0 \\ y_k = h_2(x_k) + k_{21}(x_k)w_k, \end{cases} \tag{5.49}$$

where $x \in \mathcal{X}$ is the semistate vector; $w \in \mathcal{W} \subset \Re^m$ is an unknown disturbance (or noise) signal, which belongs to the set $\mathcal{W}$ of admissible exogenous inputs; $y \in \mathcal{Y} \subset \Re^m$ is the measured output (or observation) of the system, and belongs to $\mathcal{Y}$, the set of admissible measured-outputs.

The functions $f : \mathcal{X} \to \mathcal{X}$, $g_1 : \mathcal{X} \to \mathcal{M}^{n\times m}(\mathcal{X})$, where $\mathcal{M}^{i\times j}$ is the ring of $i \times j$ smooth matrices over $\mathcal{X}$, $h_2 : \mathcal{X} \to \Re^m$, and $k_{21} : \mathcal{X} \to \mathcal{M}^{m\times m}(\mathcal{X})$ are real C^∞ functions of x, while $E \in \Re^{n\times n}$ is a constant but singular matrix. Furthermore, we assume without any loss of generality that the system (5.49) has an isolated equilibrium-point at $x = 0$ such that $f(0) = 0$, $h_2(0) = 0$. We also assume that there exists at least one solution $x(k, k_0, Ex^0, w)$ $\forall k \in \mathbf{Z}$ for the system, for all admissible initial conditions Ex^0, for all $w \in \mathcal{W}$. The initial condition Ex^0 is said to be admissible if the solution x_k is unique, impulse-free and smooth for all $k \in [k_0, \infty)$.

In addition, the following standing assumptions will be made on the system.

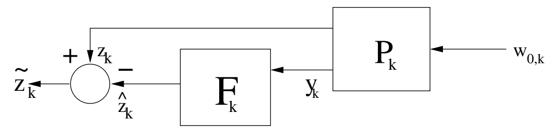

Figure 5.4 Set-up for $\mathcal{H}_2$ Filtering

Assumption 5.2.1. *Let $A = \frac{\partial f}{\partial x}(x_0)$, $x_0 \in \mathcal{O} \subset \mathcal{X}$, then the system (5.49) is admissible implies the following hold:*

1. *the system is locally regular at each point $x_0 \in \mathcal{O}$ and hence solvable, i.e, $det(zE - A) \not\equiv 0$ for all $z \in \mathbf{C}$;*
2. *the system is locally impulse-free at each $x_0 \in \mathcal{O}$, i.e., $deg(det(zE - A)) = rank(E)\ \forall z \in \mathbf{C}$;*
3. *the system is locally asymptotically stable, i.e., (E, A) is Hurwitz at $x_0 = 0$.*

The standard $\mathcal{H}_2$ local filtering/state-estimation problem is defined as follows.

Definition 5.2.1. *(Standard $\mathcal{H}_2$ Local State Estimation or Filtering Problem). Find a filter, $\mathbf{F}$, for estimating the state $x(t)$ or a function of it, $z_k = h_1(x_k)$, from observations $\mathsf{Y}_k \triangleq \{y(i) : i \leq k\}$, of $y(i)$ up to time k, to obtain the estimate*

$$\hat{x}_k = \mathbf{F}(\mathsf{Y}_k),$$

such that, the $\mathcal{H}_2$-norm from the input w to some suitable penalty function $\tilde{z}$ is locally minimized for all admissible initial conditions $Ex^0 \in \mathcal{O} \subset \mathcal{X}$. Moreover, if the filter solves the problem for all admissible $Ex^0 \in \mathcal{X}$, we say the problem is solved globally.

We shall adopt the following definition of local zero-input observability which we coined from (Ozcaldiran, 1992), (Vidyasagar, 1993).

Definition 5.2.2. *For the nonlinear system $\mathbf{P}_D^{da}$, we say that, it is locally weakly zero-input* *observable, if for all states* $x_1,\ x_2 \in U \subset \mathcal{X}$ *and input* $w(.) = 0,\ k > k_0$

$$y(k; Ex_1(k_0-), w) \equiv y(k; Ex_2(k_0-), w) \Longrightarrow Ex_1(k_0) = Ex_2(k_0); \tag{5.50}$$

the system is said to be locally zero-input observable if

$$y(k; Ex_1(k_0-), w) \equiv y(k; Ex_2(k_0-), w) \Longrightarrow x_1(k_0) = x_2(k_0); \tag{5.51}$$

where $y(., Ex_i(k_0-), w),\ \ i = 1, 2$ *is the output of the system with the initial condition* $Ex_i(t_0-)$*; and the system is said to be locally strongly zero-input observable if*

$$y(k; Ex_1(k_0-), w) \equiv y(k; Ex_2(k_0-), w) \Longrightarrow x_1(k_0-) = x_2(t_0-). \tag{5.52}$$

Moreover, the system is said to be globally (weakly, strongly) zero-input observable, if it is locally (weakly, strongly)-observable at each $x(k_0) \in \mathcal{X}$ *or* $U = \mathcal{X}$*.*

In the sequel, we shall not distinguish between local observability and strong local observability.

5.2.2 Solution to the $\mathcal{H}_2$ Filtering Problem Using Singular Filters

In this section, we discuss singular filters for the $\mathcal{H}_2$ state estimation problem defined in the previous section, and we discuss normal filters in the next subsection. For this purpose, we assume that the noise signal $w \in \mathcal{W} \subset \mathcal{S}'$ is a zero-mean Gaussian white-noise vector process with

$$\mathsf{E}\{w(k)\} = 0, \quad \mathsf{E}\{w(k)w^T(j)\} = W\delta(k-j), \quad i, j, k \in \mathbf{Z}.$$

The system's initial condition is also assumed to be Gaussian distributed random vector with mean

$$\mathsf{E}\{x_0\} = \bar{x}^0.$$

We consider full-order $\mathcal{H}_2$ singular filters for the system with the certainty-equivalent optimal noise $w^\star = \mathsf{E}w = 0$ in the usual Kalman-Luenberger type structure:

$$\mathbf{F}^{ad}_{DS1}\begin{cases} E\hat{x}_{k+1} &= f(\hat{x}_k) + L(\hat{x}_k, y_k)(y_k - h_2(\hat{x}_k)), \quad \hat{x}(k_0) = \bar{x}^0 \\ \tilde{z}_k &= y_k - h_2(\hat{x}_k), \end{cases} \tag{5.53}$$

where $\hat{x} \in \mathcal{X}$ is the filter state and $\hat{L} : \mathcal{X} \times \mathcal{Y} \to \Re^{n\times m}$ is the gain matrix of the filter, $\tilde{z} \in \Re^m$ is the penalty variable or estimation error.

The problem can then be formulated as a dynamic optimization problem with the following cost functional

$$\begin{aligned} \min_{\hat{L}\in\Re^{n\times m}, w\in\mathcal{S}'} J(L,w) &= \mathsf{E}\left\{\frac{1}{2}\sum_{k=k_0}^{\infty}\|\tilde{z}_k\|^2_W\right\} = \frac{1}{2}\left\{\|\mathbf{F}^{da}_S \circ \mathbf{P}^{da}_D\|^2_{\mathcal{H}_2}\right\}_W, \\ & s.t.\ (5.53), \quad \text{and with } w = 0, \quad \lim_{t\to\infty}\{\hat{x} - x\} = 0; \end{aligned} \tag{5.54}$$

To solve the above problem, we form the Hamiltonian function[2] $H : \mathcal{X} \times \mathcal{Y} \times \Re^{n\times m} \times \Re \to \Re$ (Wang, 2008):

$$H(\hat{x}, y, L, V) = V(E(f(\hat{x}) + L(\hat{x}, y)(y - h_2(\hat{x}))), y) - V(E\hat{x}, y_{k-1}) + \frac{1}{2}\|\tilde{z}\|^2_W \tag{5.55}$$

for some C^2 function $V : \mathcal{X} \times \mathcal{Y} \to \Re$ and where $x = x_k$, $y = y_k$, $\tilde{z} = \tilde{z}_k$. Notice also here that, we are only using y_{k-1} in the above expression (5.55) to distinguish between $y_k = y$ and y_{k-1}. Otherwise, (5.55) holds for all y and is smooth in all its arguments.

Then, the optimal gain $L^\star$ can be obtained by minimizing H with respect to L in the above expression (5.55), as

$$L^\star = \arg\min_{\hat{L}} H(\hat{x}, y, L, V). \tag{5.56}$$

Because the Hamiltonian function (5.55) is not a linear or quadratic function of the gain L,

[2]Our definition is slightly different from Reference (Wang, 2008) in order to maintain the symmetry of the Hamiltonian

only implicit solutions can be obtained by solving the equations

$$\left.\frac{\partial V(\lambda, y)}{\partial \lambda}\right|_{\lambda=\lambda^\star} = 0 \tag{5.57}$$

for $L^\star(\hat{x}, y)$, where $\lambda = E(f(\hat{x}) + L(\hat{x}, y)(y - \tilde{h}_2(\hat{x})))$, $\partial V/\partial\lambda$ is the row vector of first-order partial derivatives of V with respect to λ, and V solves the discrete-time Hamilton-Jacobi-Bellman equation (DHJBE)

$$H(\hat{x}, y, L^\star, V) = 0, \quad V(0,0) = 0, \tag{5.58}$$

with

$$\left.\frac{\partial^2 V}{\partial \lambda^2}\right|_{\lambda=\lambda^\star} > 0.$$

Thus, the only way to obtain an explicit solution is to use an approximate scheme. Accordingly, consider a second-order quadratic approximation of the Hamiltonian function (5.55) about $(Ef(\hat{x}), y)$ and in the direction of the estimator state vector $E\hat{x}$, denoted by $\widehat{H}$:

$$\begin{aligned}\widehat{H}(\hat{x}, y, \hat{L}, \hat{V}) \;=\; & \hat{V}(Ef(\hat{x}), y) + \hat{V}_{E\hat{x}}(Ef(\hat{x}), y)[E(f(\hat{x}) + \hat{L}(\hat{x}, y)(y - h_2(\hat{x})))] + \\ & \frac{1}{2}(y - h_2(\hat{x}))^T \hat{L}^T(\hat{x}, y) E^T \hat{V}_{E\hat{x}E\hat{x}}(Ef(\hat{x}), y) E\hat{L}(\hat{x}, y)(y - h_2(\hat{x})) - \\ & \hat{V}(E\hat{x}, y_{k-1}) + \frac{1}{2}\|\tilde{z}\|^2_W + O(\|\hat{x}\|^3),\end{aligned} \tag{5.59}$$

where $\hat{V}$, $\hat{L}$, are the corresponding approximate functions with $\hat{V}$ positive-definite, and $\hat{V}_{E\hat{x}E\hat{x}}$ is the Hessian matrix of $\hat{V}$ with respect to $E\hat{x}$. Then, differentiating $\widehat{H}(., ., \hat{L}, .)$ with respect to $u = E\hat{L}(\hat{x}, y)(y - h_2(\hat{x}))$ and applying the necessary optimality conditions, i.e., $\frac{\partial \widehat{H}}{\partial u} = 0$, we get

$$E\hat{L}^\star(\hat{x}, y)(y - h_2(\hat{x})) = -[\hat{V}_{E\hat{x}E\hat{x}}(Ef(\hat{x}), y)]^{-1} \hat{V}^T_{E\hat{x}}(Ef(\hat{x}), y). \tag{5.60}$$

Finally, substituting the above expression for $\hat{L}^\star$ in (5.59) and setting

$$\widehat{H}(\hat{x}, y, \hat{L}^\star, \hat{V}) = 0,$$

results in the following DHJBE:

$$
\begin{aligned}
&\hat{V}(Ef(\hat{x}),y)+\hat{V}_{E\hat{x}}(Ef(\hat{x}),y)Ef(\hat{x})-\hat{V}(E\hat{x},y_{k-1})- \\
&\tfrac{1}{2}\hat{V}_{E\hat{x}}(Ef(\hat{x}),y)[\hat{V}_{E\hat{x}E\hat{x}}(Ef(\hat{x}),y)]^{-1}\hat{V}^T_{E\hat{x}}(Ef(\hat{x}),y)+ \\
&\tfrac{1}{2}(y-h_2(\hat{x}))^TW(y-h_2(\hat{x}))=0, \quad \hat{V}(0,0)=0.
\end{aligned}
\tag{5.61}
$$

We then have the following result.

Theorem 5.2.1. *Consider the nonlinear system (5.49) and the $\mathcal{H}_2$ filtering problem for this system. Suppose the plant $\mathbf{P}^{\mathbf{ad}}_{\mathbf{D}}$ is locally asymptotically stable about the equilibrium-point $x=0$ and zero-input observable. Further, suppose there exist a C^2 positive-definite function $\hat{V}:\hat{N}\times\hat{\Upsilon}\to\Re_+$ locally defined in a neighborhood $\hat{N}\times\hat{\Upsilon}\subset\mathcal{X}\times\mathcal{Y}$ of the origin $(\hat{x},y)=(0,0)$, and a matrix function $\hat{L}:\hat{N}\times\hat{\Upsilon}\to\Re^{n\times m}$, satisfying the DHJBE (5.61) together with the side-condition (5.60). Then, the filter $\mathbf{F}^{ad}_{DS1}$ solves the $\mathcal{H}_2$ filtering problem for the system locally in $\hat{N}$.*

Proof: The optimality of the filter gain $\hat{L}^\star$ has already been shown above. It remains to prove asymptotic convergence of the estimation error vector. Accordingly, let $\hat{V}(E\hat{x},y)\geq 0$ be a C^1 solution of the DHJBE (5.61). Then, taking the time variation of $\hat{V}$ along a trajectory of (5.53), with $\hat{L}=\hat{L}^\star$, we get

$$
\begin{aligned}
\hat{V}(E\hat{x}_{k+1},y_k) \;\approx\;& \hat{V}(Ef(\hat{x}),y)+\hat{V}_{E\hat{x}}(Ef(\hat{x}),y)[E(f(\hat{x})+\hat{L}(\hat{x},y)(y-h_2(\hat{x})))]+ \\
&\frac{1}{2}(y-h_2(\hat{x}))^T\hat{L}^T(\hat{x},y)E^T\hat{V}_{E\hat{x}E\hat{x}}(Ef(\hat{x}),y)E\hat{L}(\hat{x},y)(y-h_2(\hat{x})) \\
=\;& \hat{V}(Ef(\hat{x}),y)+\hat{V}_{E\hat{x}}(Ef(\hat{x}),y)Ef(\hat{x})- \\
&\frac{1}{2}\hat{V}_{E\hat{x}}(Ef(\hat{x}),y)[\hat{V}_{E\hat{x}E\hat{x}}(Ef(\hat{x}),y)]^{-1}\hat{V}^T_{E\hat{x}}(Ef(\hat{x}),y) \\
=\;& \hat{V}(E\hat{x},y_{k-1})-\frac{1}{2}(y-h_2(\hat{x}))^TW(y-h_2(\hat{x}))
\end{aligned}
$$

where use has been made of the quadratic Taylor approximation above, and the last equality follows from using the DHJBE (5.61). Therefore,

$$\hat{V}(E\hat{x}_{k+1}, y_k) - \hat{V}(E\hat{x}, y_{k-1}) = -\frac{1}{2}\|\tilde{z}_k\|^2_W$$

and by Lyapunov's theorem, the filter dynamics is stable, i.e., $\hat{V}(E\hat{x}, y)$ is non-increasing along a trajectory of (5.53). Further, the condition that $\hat{V}(E\hat{x}_{k+1}, y_k) \equiv \hat{V}(E\hat{x}, y_{k-1})$ $\forall k \geq k_s$, for some k_s, implies that $\tilde{z}_k \equiv 0$, which further implies that $y_k = h_2(\hat{x}_k)$ $\forall k \geq k_s$. By the zero-input observability of the system, this implies that $\hat{x}_k = x_k$ $\forall k \geq k_s$. □

The result of the theorem can be specialized to the linear descriptor system

$$\mathbf{P}^{dl}_D : \begin{cases} Ex_{k+1} &= Ax_k + B_1 w_k; \quad Ex(k_0) = Ex^0 \\ y_k &= C_2 x_k + D_{21} w_k \end{cases} \tag{5.62}$$

where $E \in \Re^{n\times n}$, $A \in \Re^{n\times n}$, $B_1 \in \Re^{n\times m}$, $C_2 \in \Re^{m\times n}$, $D_{21} \in \Re^{m\times m}$. Assuming without loss of generality that $W = I$, we have the following result.

Corollary 5.2.1. *Consider the linear descriptor system (5.62) and the $\mathcal{H}_2$ filtering problem for this system. Suppose the plant $\mathbf{P}^{dl}_D$ is locally asymptotically stable about the equilibrium-point $x = 0$ and observable. Suppose further, there exist symmetric positive-semidefinite matrices $\hat{P} \in \Re^{n\times n}$, $\hat{Q}, \hat{R} \in \Re^{m\times m}$, and a matrix $\hat{L} \in \Re^{n\times m}$, satisfying the linear matrix-inequalities (LMIs)*

$$\begin{bmatrix} A^T E^T \hat{P} E A + E^T \hat{P} E + \frac{1}{2} C_2^T C_2 & -\frac{1}{2} C_2^T & 0 \\ -\frac{1}{2} C_2 & -\frac{1}{2}(\hat{Q} - \hat{R}) & 0 \\ 0 & 0 & -\frac{1}{2}\hat{Q} \end{bmatrix} \leq 0 \tag{5.63}$$

$$\begin{bmatrix} E(A - \hat{L}C_2) & \frac{1}{2} E\hat{L} \\ \frac{1}{2}\hat{L}^T E^T & -\delta_1 I \end{bmatrix} \leq 0 \tag{5.64}$$

for some number $\delta_1 \geq 0$. Then the filter

$$\mathbf{F}^{dl}_{DS1} : \begin{cases} E\dot{\hat{x}} &= A\hat{x} + \hat{L}(y - C_2\hat{x}); \quad E\hat{x}(k_0) = E\bar{x}^0 \\ \hat{z} &= C_2\hat{x} \end{cases} \tag{5.65}$$

solves the $\mathcal{H}_2$ estimation problem for the system.

Proof: Take

$$\hat{V}(E\hat{x}, y) = \frac{1}{2}(\hat{x}^T E^T \hat{P} E\hat{x} + y^T \hat{Q} y), \quad \hat{P} > 0$$

and apply the result of the Theorem. □

Notice that the DHJIE (5.61) is a second-order PDE, and hence there is an increased computational burden in finding its solution. Thus, alternatively, the results of Proposition 5.2.1 can be rederived using a first-order Taylor-series approximation of the Hamiltonian (5.55), which can be obtained from (5.59) by neglecting the quadratic term, as

$$\begin{aligned}\widehat{H}_1(\hat{x}, y, \hat{L}, \hat{V}) &= \hat{V}(Ef(\hat{x}), y) + \hat{V}_{E\hat{x}}(Ef(\hat{x}), y)[E(f(\hat{x}) + \hat{L}(\hat{x}, y)(y - h_2(\hat{x})))] - \\ &\quad \hat{V}(E\hat{x}, y_{k-1}) + \frac{1}{2}\|\tilde{z}\|^2 + O(\|\hat{x}\|^2). \end{aligned} \tag{5.66}$$

Then, repeating the optimization as previously, we can arrive at the following first-order counterpart of Proposition 5.2.1

Proposition 5.2.1. *Consider the nonlinear system (5.49) and the $\mathcal{H}_2$ filtering problem for this system. Suppose the plant $\mathbf{P}^{\mathbf{ad}}_{\mathbf{D}}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and zero-input observable. Further, suppose there exist a C^1 positive-semidefinite function $\hat{Y} : \hat{N}_1 \times \hat{\Upsilon}_1 \to \Re_+$ locally defined in a neighborhood $\hat{N}_1 \times \hat{\Upsilon}_1 \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\hat{x}, y) = (0, 0)$, and a matrix function $\hat{L} : \breve{N}_1 \times \breve{\Upsilon}_1 \to \Re^{n\times m}$, satisfying the DHJBE (Lyapunov equation)*

$$\begin{aligned}&\hat{Y}(Ef(\hat{x}), y) + \hat{Y}_{E\hat{x}}(Ef(\hat{x}), y)Ef(\hat{x}) - \hat{Y}(E\hat{x}, y_{k-1}) + \\ &\frac{1}{2}(y - h_2(\hat{x}))^T(W - 2I)(y - h_2(\hat{x})) = 0, \quad \hat{Y}(0, 0) = 0,\end{aligned} \tag{5.67}$$

together with the side-condition

$$\hat{Y}_{E\hat{x}}(Ef(\hat{x}), y)E\hat{L}^\star(\hat{x}, y) = -(y - h_2(\hat{x}))^T. \tag{5.68}$$

Then, the filter $\mathbf{F}^{ad}_{DS1}$ solves the $\mathcal{H}_2$ filtering problem for the system locally in $\hat{N}_1$.

Proof: Proof follows along same lines as Theorem 5.2.1. □

In the next section, we consider the design of normal filters for the system.

5.2.3 Discrete-time $\mathcal{H}_2$ Normal Filters

In this subsection, we discuss normal filters for the system (5.49). We shall consider the design of both full-order and reduced-order filters. We start with the full-order filter first, and in this regard, without any loss of generality, we can assume that E is of the form

$$E = \begin{pmatrix} I_{q\times q} & 0 \\ 0 & 0 \end{pmatrix}.$$

This follows from matrix theory and can easily be proven using the singular-value decomposition (SVD) of E. It follows that, the system can be represented in the canonical form of a differential-algebraic system

$$\bar{\mathbf{P}}_D^{ad} : \begin{cases} x_{1,k+1} &= f_1(x_k) + g_{11}(x_k)w_k; \;\; x(k_0) = x^0 \\ 0 &= f_2(x_k) + g_{21}(x_k)w_k \\ y &= h_2(x_k) + k_{21}(x_k)w_k, \end{cases} \tag{5.69}$$

where $dim(x_1) = q$, $f_1(0) = 0$, $f_2(0) = 0$. Then, if we define

$$x_{2,k+1} = f_2(x_k) + g_{21}(x_k)w_k,$$

where $x_{2,k+1}$ is a fictitious state vector, and $dim(x_2) = n - q$, the system (5.69) can be represented by a normal state-space system as

$$\widetilde{\mathbf{P}}_D^{ad} : \begin{cases} x_{1,k+1} &= f_1(x_k) + g_{11}(x_k)w_k; \;\; x_1(k_0) = x^{10} \\ x_{2,k+1} &= f_2(x_k) + g_{21}(x_k)w_k; \;\; x_2(k_0) = x^{20} \\ y &= h_2(x_k) + k_{21}(x_k)w_k. \end{cases} \tag{5.70}$$

Now define the following set $\Omega_o \subset \mathcal{X}$

$$\Omega_o = \{(x_1, x_2) \in \mathcal{X} \mid x_{2,k+1} \equiv 0\}. \tag{5.71}$$

Then, we have the following system equivalence

$$\widetilde{\mathbf{P}}^{ad}_{D}|_{\Omega_o} = \bar{\mathbf{P}}^{ad}_{D}. \tag{5.72}$$

Therefore, to estimate the states of the system (5.69), we need to stabilize the system (5.70) about Ω_o, and then design a filter for the resulting system. For this purpose, we consider the following class of filters with $\grave{w}^\star = \mathsf{E}\{w\} = 0$

$$\mathbf{F}^{ad}_{DN3} \begin{cases} \grave{x}_{1,k+1} &= f_1(\grave{x}_k) + \grave{L}_1(\grave{x}_k, y_k)(y_k - h_2(\grave{x}_k)) \\ \grave{x}_{2,k+1} &= f_2(\grave{x}_k) + g_{22}(x_k)\alpha_2(\grave{x}_k) + \grave{L}_2(\grave{x}_k, y_k)(y_k - h_2(\grave{x}_k)) \\ \grave{z}_k &= y_k - h_2(\grave{x}_k), \end{cases} \tag{5.73}$$

where $\grave{x} \in \mathcal{X}$ is the filter state, $\grave{L}_1 : \mathcal{X} \times \mathcal{Y} \to \Re^{q \times m}$, $\grave{L}_2 : \mathcal{X} \times \mathcal{Y} \to \Re^{n-q \times m}$ are the filter gain matrices, and $\tilde{g}_{22} : \mathcal{X} \to \mathcal{M}^{(n-q)\times p}$ is a gain matrix for the artificial control input $u = \alpha_2(\grave{x}) \in \Re^p$ required to stabilize the dynamics $x_{2,k+1}$ about Ω_o. Accordingly, we make the following assumption.

Assumption 5.2.2. *The pair $\{f_2, \tilde{g}_{22}\}$ is locally stabilizable, i.e., $\exists$ a control law $\alpha_2(\grave{x}_2)$ and a Lyapunov-function (LF), $\bar{V} > 0$, such that $\bar{V}(f_2(\grave{x}) - \tilde{g}_{22}(\grave{x})\alpha_2(\grave{x})) - \bar{V}(\grave{x}) < 0 \;\forall \grave{x} \in \grave{N} \subset \mathcal{X}$.*

Thus, if Assumption 5.2.2 holds, then we can make $\alpha_2 = \alpha_2(\grave{x}, \varepsilon)$, where $\varepsilon > 0$ is small, a high-gain feedback (Young, 1977) to constrain the dynamics on Ω_o as fast as possible. Then, we proceed to design the gain matrices $\grave{L}_1$, $\grave{L}_2$ to estimate the states using similar approximations as in the previous section. Using the first-order Taylor approximation, we have the following result.

Proposition 5.2.2. *Consider the nonlinear system (5.69) and the $\mathcal{H}_2$ estimation problem for this system. Suppose the plant $\bar{\mathbf{P}}^{\mathbf{ad}}_{\mathbf{D}}$ is locally asymptotically stable about the equilibrium-point $x = 0$, and zero-input observable. Further, suppose there exist a C^1 positive-semidefinite* *function* $\grave{V} : \grave{N} \times \grave{\Upsilon} \to \Re_+$*, locally defined in a neighborhood* $\grave{N} \times \grave{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ *of the origin* $(\grave{x}_1, \grave{x}_2, y) = (0, 0, 0)$*, and matrix functions* $\grave{L}_1 : \grave{N} \times \grave{\Upsilon} \to \Re^{q \times m}$*,* $\grave{L}_2 : \grave{N} \times \grave{\Upsilon} \to \Re^{n-q \times m}$*, satisfying the DHJBE:*

$$\begin{aligned} &\grave{V}(f_1(\grave{x}), f_2(\grave{x}), y) - \grave{V}(\grave{x}_1, \grave{x}_2, y_{k-1}) + \grave{V}_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y) g_{22}(\grave{x}) \alpha_2(\grave{x}, \varepsilon) + \\ &\quad \frac{1}{2}(y - h_2(\grave{x}))^T (W - 4I)(y - h_2(\grave{x})) = 0, \quad \grave{V}(0,0,0) = 0, \end{aligned} \tag{5.74}$$

together with the side-conditions

$$\grave{V}_{\grave{x}_1}(f_1(\grave{x}), f_2(\grave{x}), y)\grave{L}_1(\grave{x}_1, \grave{x}_2, y) = -(y - h_2(\grave{x}))^T, \tag{5.75}$$

$$\grave{V}_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y)\grave{L}_2(\grave{x}_1, \grave{x}_2, y) = -(y - h_2(\grave{x}))^T. \tag{5.76}$$

Then, the filter $\mathbf{F}^{da}_{DN3}$ *solves the* $\mathcal{H}_2$*-filtering problem for the system locally in* $\grave{N}$*.*

Proof: Follows along same lines as Theorem 5.2.1.

A common DHJBE for both the stabilization and filter design can also be utilized in the above design procedure. This can be achieved optimally if we take

$$\begin{aligned} \alpha_2(\grave{x}, \varepsilon) &= -\frac{1}{\varepsilon} g_{22}^T(\grave{x}) \bar{V}^T_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y), \\ \bar{V}_{\grave{x}_1}(f_1(\grave{x}), f_2(\grave{x}), y)\grave{L}_1(\grave{x}, y) &= -(y - h_2(\grave{x}))^T, \\ \bar{V}_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y)\grave{L}_2(\grave{x}, y) &= -(y - h_2(\grave{x}))^T, \end{aligned}$$

where $\bar{V} \geq 0$ is a C^1 solution of the following DHJBE

$$\begin{aligned} &\bar{V}(f_1(\grave{x}), f_2(\grave{x}), y) - \frac{1}{\varepsilon}\bar{V}_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y) g_{22}(\grave{x}) g_{22}^T(\grave{x}) \bar{V}^T_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y) - \\ &\bar{V}(\grave{x}_1, \grave{x}_2, y_{k-1}) + \frac{1}{2}(y - h_2(\grave{x}))^T (W - 4I)(y - h_2(\grave{x})) = 0, \quad \bar{V}(0,0,0) = 0. \end{aligned}$$

Next, we consider a reduced-order normal filter design. Accordingly, partition the state-vector x comformably with $rank(E) = q$ as $x = (x_1^T \;\; x_2^T)^T$ with $dim(x_1) = q$, $dim(x_2) = n-q$

and the state equations as

$$\breve{\mathbf{P}}^{ad}_{D} : \begin{cases} x_{1,k+1} &= f_1(x_{1,k}, x_{2,k}) + g_{11}(x_{1,k}, x_{2,k})w_k; \quad x_1(k_0) = x^{10} \\ 0 &= f_2(x_{1,k}, x_{2,k}) + g_{21}(x_{1,k}, x_{2,k})w_k; \quad x_2(k_0) = x^{20} \\ y_k &= h_2(x_k) + k_{21}(x_k)w_k. \end{cases} \tag{5.77}$$

Then we make the following assumption.

Assumption 5.2.3. *The system is in the standard-form, i.e., the Jacobian matrix* $f_{2,x_2}(x_1, x_2)$ *is nonsigular in an open neighborhood* $\widetilde{U}$ *of* $(0, 0)$ *and* $g_{21}(0, 0) \neq 0$.

If Assumption 5.2.3 holds, then by the Implicit-function Theorem (Sastry, 1999), there exists a unique C^1 function $\phi : \Re^q \times \mathcal{W} \to \Re^{n-q}$ and a solution

$$\bar{x}_2 = \phi(x_1, w)$$

to equation (5.77b). Thus, the system can be locally represented in $\widetilde{U}$ as the reduced-order system

$$\bar{\mathbf{P}}^{ad}_{rD} : \begin{cases} x_{1,k+1} &= f_1(x_{1,k}, \phi(x_{1,k}, w_k)) + g_{11}(x_{1,k}, \phi(x_{1,k}, w_k))w_k; \quad x_1(k_0) = x^{10} \\ y_k &= h_2(x_{1,k}, \phi(x_{1,k}, w_k)) + k_{21}(x_{1,k}, \phi(x_{1,k}, w_k))w_k. \end{cases} \tag{5.78}$$

We can then design a normal filter of the form

$$\mathbf{F}^{ad}_{DrN4} \begin{cases} \breve{x}_{1,k+1} &= f_1(\breve{x}_{1,k}, \phi(\breve{x}_{1,k}, 0)) + \breve{L}(\breve{x}_{1,k}, \phi(\breve{x}_{1,k}, 0), y_k)[y- \\ & \quad h_2(\breve{x}_{1,k}, \phi(\breve{x}_{1,k}, 0))]; \quad \breve{x}_1(k_0) = \bar{x}^{10} \\ \breve{z}_k &= y_k - h_2(\breve{x}_{1,k}, \phi(\breve{x}_{1,k}, 0)) \end{cases} \tag{5.79}$$

for the system, and consequently, we have the following result.

Theorem 5.2.2. *Consider the nonlinear system (5.69) and the* $\mathcal{H}_2$ *filtering problem for this system. Suppose the plant* $\bar{\mathbf{P}}^{\mathbf{ad}}_{\mathbf{D}}$ *is locally asymptotically stable about the equilibrium-point* $x = 0$, *zero-input observable and Assumption 5.2.3 holds for the system. Further, suppose there exists a* C^1 *positive-semidefinite function* $\breve{V} : \breve{N} \times \breve{\Upsilon} \to \Re_+$, *locally defined* *in a neighborhood* $\breve{N} \times \breve{\Upsilon} \subset \widetilde{U} \times \mathcal{Y}$ *of the origin* $(\breve{x}_1, y) = (0, 0)$*, and a matrix function* $\breve{L} : \breve{N} \times \breve{\Upsilon} \to \Re^{q \times m}$*, satisfying the DHJBE:*

$$\begin{aligned} &\breve{V}(f_1(\breve{x}_1, \phi(\breve{x}_1, 0), y) + \breve{V}_{\breve{x}_1}(f_1(\breve{x}_1, \phi(\breve{x}_1, 0), y) f_1(\breve{x}_1, \phi(\breve{x}_1, 0)) - \breve{V}(\breve{x}_1, y_{k-1}) + \\ &\frac{1}{2}(y - h_2(\breve{x}_1, \phi(\breve{x}_1, 0)))^T (W - 2I)(y - h_2(\breve{x}_1, \phi(\breve{x}_1, 0))) = 0,\ \breve{V}(0, 0) = 0, \end{aligned} \tag{5.80}$$

together with the side-condition

$$\breve{V}_{\breve{x}_1}(f_1(\breve{x}_1, \phi(\breve{x}_1, 0), y)\breve{L}(\breve{x}_1, y) = -(y - h_2(\breve{x}_1, \phi(\breve{x}_1, 0)))^T. \tag{5.81}$$

Then, the filter $\mathbf{F}^{ad}_{DrN4}$ *solves the* $\mathcal{H}_2$ *local filtering problem for the system in* $\breve{N}$*.*

Proof: Follows along same lines as Theorem 5.2.2.

Similarly, we can specialize the result of Theorem 5.2.2 to the linear system (5.62). The system can be rewritten in the form (5.69) as

$$\mathbf{P}^{ld}_D \ : \begin{cases} x_{k+1} &= A_1 x_{1,k} + A_{12} x_{2,k} + B_{11} w_k; \quad x_1(k_0) = x^{10} \\ 0 &= A_{21} x_{1,k} + A_2 x_{2,k} + B_{21} w_k; \quad x_2(k_0) = x^{20} \\ y_k &= C_{21} x_{1,k} + C_{22} x_{2,k} + D_{21} w_k. \end{cases} \tag{5.82}$$

Then, if A_2 is nonsingular (Assumption 5.2.3) we can solve for x_2 in equation (5.82(b)) to get

$$\bar{x}_2 = -A_2^{-1}(A_{21} x_1 + B_{21} w)$$

and the filter (5.79) takes the following form

$$\mathbf{F}^{ld}_{DrN4} \begin{cases} \breve{x}_{1,k+1} &= (A_1 - A_2^{-1} A_{21})\breve{x}_{1,k} + \breve{L}(y_k - (C_{21} - C_{22} A_2^{-1} A_{21})\breve{x}_{1,k}); \quad \breve{x}_1(k_0) = \bar{x}^{10} \\ \breve{z}_k &= y_k - (C_{21} - C_{22} A_2^{-1} A_{21})\breve{x}_{1,k}. \end{cases} \tag{5.83}$$

Moreover, if we assume without loss of generality $W = I$, we have the following corollary.

Corollary 5.2.2. *Consider the linear descriptor system (5.62) and the* $\mathcal{H}_2$*-filtering problem* *for this system. Suppose the plant* $\mathbf{P}_D^{ld}$ *is locally asymptotically stable about the equilibrium-point* $x = 0$, *Assumption 5.2.3 holds, and the plant is observable. Suppose further, there exist symmetric positive-semidefinite matrices* $\breve{P} \in \Re^{q\times q}$, $\breve{Q} \in \Re^{m\times m}$, *and a matrix* $\breve{L} \in \Re^{n\times m}$, *satisfying the LMIs:*

$$\begin{bmatrix} 3\tilde{A}_1^T\breve{P}\tilde{A}_1 - \breve{P} - \tilde{C}_2^T\tilde{C}_2^T & \tilde{C}_2^T & 0 \\ \tilde{C}_2 & \breve{Q} - I & 0 \\ 0 & 0 & -\breve{Q} \end{bmatrix} \leq 0 \tag{5.84}$$

$$\begin{bmatrix} 0 & \frac{1}{2}(\tilde{A}_1^T\breve{P}\breve{L} - \tilde{C}_2^T) \\ \frac{1}{2}(\tilde{A}_1^T\breve{P}\breve{L} - \tilde{C}_2^T)^T & (1-\delta_3)I \end{bmatrix} \leq 0 \tag{5.85}$$

for some number $\delta_3 \geq 1$, *where* $\tilde{A}_1 = (A_1 - A_2^{-1}A_{21})$, $\tilde{C}_2 = (C_{21} - C_{22}A_2^{-1}A_{21})$. *Then, the filter (5.83) solves the* $\mathcal{H}_2$*-filtering problem for the system.*

Proof: Take

$$\breve{V}(\breve{x}) = \frac{1}{2}(\breve{x}_1^T\breve{P}\breve{x}_1 + y^T\breve{Q}y)$$

and apply the result of the Theorem. □

5.2.4 The General Discrete-time case

In this section, we consider the filtering problem for the more general class of affine descriptor systems in which $E = E(x) \in \mathcal{M}^{n\times n}(\mathcal{X})$ is a matrix function of x, and can be represented as

$$\mathbf{P}_{DG}^{ad} : \begin{cases} E(x_k)x_{k+1} &= f(x_k) + g_1(x_k)w_k; \quad x(k_0) = x^0 \\ y_k &= h_2(x_k) + k_{21}(x_k)w_k \end{cases} \tag{5.86}$$

where *minimum rank*$(E(x)) = q$ for all $x \in \mathcal{X}$, $E(0) = 0$, and all the other variables and functions have their previous meanings and dimensions.

We first consider the design of a singular filter for the above system. Accordingly, consider a filter of the form (5.53) for the system defined as

$$\mathbf{F}^{ad}_{DS5} \begin{cases} E(\check{x}_k)\check{x}_{k+1} &= f(\check{x}_k) + \check{L}(\check{x}_k, y_k)(y_k - h_2(\check{x}_k)) \\ \check{z}_k &= y_k - h_2(\check{x}_k), \end{cases} \tag{5.87}$$

where $\check{L} \in \Re^{n\times m}$ is the gain of the filter. Suppose also the following assumption holds.

Assumption 5.2.4. *There exists a vector-field $e(x) = (e_1(x), \ldots, e_n(x))^T$ such that*

$$E(x) = \frac{\partial e}{\partial x}(x), \quad e(0) = 0.$$

Remark 5.2.1. *Notice that, $e(x)$ cannot in general be obtained by line-integration of the rows of $E(x)$.*

Then we have the following result.

Theorem 5.2.3. *Consider the nonlinear system (5.78) and the $\mathcal{H}_2$ state estimation problem for this system. Suppose for the plant $\mathbf{P^{ad}}_{DG}$ is locally asymptotically stable about the equilibrium-point $x = 0$, and zero-input observable. Further, suppose Assumption 5.2.4 holds, there exists a C^1 positive-semidefinite function $\check{Y} : \check{N}\times\check{\Upsilon} \to \Re_+$, locally defined in a neighborhood $\check{N}\times\check{\Upsilon} \subset \mathcal{X}\times\mathcal{Y}$ of the origin $(e(x), y) = (0, 0)$, and a matrix function $\check{L} : \check{N}\times\check{\Upsilon} \to \Re^{n\times m}$, satisfying the DHJBE:*

$$\begin{aligned} &\check{Y}(E(\check{x})f(\check{x}), y) + \check{Y}_{e(\check{x})}(E(x)f(\check{x}), y)E(\check{x})f(\check{x}) - \check{Y}(e(\check{x}), y_{k-1}) + \\ &\frac{1}{2}(y - h_2(\check{x}))^T(W - 2I)(y - h_2(\check{x})) = 0, \quad \check{Y}(0,0) = 0 \end{aligned} \tag{5.88}$$

together with the side-condition

$$\check{Y}_{e(\check{x})}(E(\check{x})f(\check{x}), y)E(\check{x})\check{L}^\star(\check{x}, y) = -(y - h_2(\check{x}))^T. \tag{5.89}$$

Then, the filter $\mathbf{F}^{ad}_{DS5}$ solves the $\mathcal{H}_2$ local filtering problem for the system in $\check{N}$.

Proof: Let $\check{Y} \geq 0$ be a C^1 solution of the DHJBE (5.80), and consider the time-variation of this function along a trajectory of (5.87) with $\check{L}(\check{x}, y) = \check{L}^\star(\check{x}, y)$

$$\begin{aligned}
\check{Y}(E(\check{x})\check{x}_{k+1}, y) &\approx \check{Y}(E(\check{x})f(\check{x}), y) + \check{Y}_{e(\check{x})}(E(x)f(\check{x}), y)E(\check{x})f(\check{x}) + \\
&\quad \check{Y}_{e(\check{x})}(E(\check{x})f(\check{x}), y)E(\check{x})\check{L}^\star(\check{x}, y)(y - h_2(\check{x})) \\
&= \check{Y}(e(\check{x}), y_{k-1}) - \frac{1}{2}\|\check{z}\|_W^2,
\end{aligned}$$

where in the above manipulations, we have used the first-order Taylor approximation, and the last equality follows from the DHJBE (5.88). Thus, again by Lyapunov's Theorem, the estimator dynamics is stable. The rest of the proof then follows along the same lines as Theorem 5.2.2. □

A normal filter for the system can also be designed. If $rank(E(x)) = q$ is constant for all $x \in \widetilde{\Upsilon} \subset \mathcal{X}$, then, it can be shown (Zimmer, 1997) that, there exists a nonsingular transformation $T : \widetilde{\Upsilon} \to \mathcal{M}^{n\times n}(\mathcal{X})$ such that

$$T(x)E(x) = \begin{pmatrix} E_1(x) \\ 0 \end{pmatrix}, \quad T(x)f(x) = \begin{pmatrix} \tilde{f}_1(x) \\ \tilde{f}_2(x) \end{pmatrix},$$

where $E_1 \in \mathcal{M}^{q\times q}(\widetilde{\Upsilon})$ is nonsingular on $\widetilde{\Upsilon}$. The system (5.86) can then be similarly represented in this coordinates as

$$\widetilde{\mathbf{P}}^{\mathbf{a}}_{\mathbf{DG}} : \begin{cases} x_{1,k+1} = E_1^{-1}(x_k)\tilde{f}_1(x_{1,k}, x_{2,k}) + E_1^{-1}(x_k)\tilde{g}_{11}(x_{1,k}, x_{2,k})w_k; \; x_1(k_0) = x^{10} \\ 0 = \tilde{f}_2(x_{1,k}, x_{2,k}) + \tilde{g}_{21}(x_{1,k}, x_{2,k})w_k; \quad x_2(k_0) = x^{20} \\ y_k = h_2(x_k) + k_{21}(x_k)w_k, \end{cases} \tag{5.90}$$

where $\begin{pmatrix} \tilde{g}_{11}(x) \\ \tilde{g}_{21}(x) \end{pmatrix} = T(x)g_1(x)$. A normal filter can now be designed for the above transformed system using the procedure outlined in Subsection 5.2.3 and Proposition 5.2.2. Similarly, a reduced-order filter for the system can also be designed as in Theorem 5.2.2 if the equivalent of Assumption 5.2.2 is satisfied for the system. This would also circumvent the problem of satisfying Assumption 5.2.4.

5.2.5 Examples

Consider the following simple nonlinear differential-algebraic system:

$$x_{1,k+1} = x_{1,k}^{1/3} + x_{2,k}^{1/5} + sin(x_{1,k})w_{0,k} \tag{5.91}$$

$$0 = x_{1,k} + x_{2,k} \tag{5.92}$$

$$y_k = x_{1,k} + x_{2,k} + w_{0,k}. \tag{5.93}$$

where w_0 is a zero-mean Gaussian white-noise process with unit variance. A singular filter of the form $\mathbf{F}^a_{DS1}$ (5.53) presented in Subsection 5.2.2 can be designed. It can be checked that, the system is locally zero-input observable, and the function $\hat{V}(\hat{x}) = \frac{1}{2}(\hat{x}_1^2 + \hat{x}_2^2 + y^2)$, solves the DHJBE (5.67) for the system. Subsequently, we calculate the gain of the filter as

$$\hat{l}_1(\hat{x}_k, y_k) = -\frac{(y_k - \hat{x}_{1,k} - \hat{x}_{2,k})}{\hat{x}_{1,k}^{1/3} + \hat{x}_{2,k}^{1/5}},$$

where $\hat{l}_1$ is set equal zero if $|\hat{x}_{1,k}^{1/3} + \hat{x}_{2,k}^{1/5}| < \epsilon$ (ϵ small) to avoid a singularity. Thus, $x_{1,k}$ can be estimated with the filter, while $x_{2,k}$ can be estimated from $\hat{x}_{2,k} = -\hat{x}_{1,k}$.

Similarly, a normal filter of the form (5.79) can be designed. It can be checked that, Assumption 5.2.3 is satisfied, and the function $\breve{V}(\breve{x}) = \frac{1}{2}(\breve{x}_1^2 + y^2)$ solves the DHJBE (5.77) for the system. Consequently, we can also calculate the filter gain as

$$\breve{l}_1(\breve{x}_k, y_k) = -\frac{(y_k - \hat{x}_{1,k} - \breve{x}_{2,k})}{\breve{x}_{1,k}^{1/3} + \breve{x}_{1,k}^{1/5}}$$

and again $\breve{l}_1$ is set equal zero if $|\breve{x}_{1,k}^{1/3} + \breve{x}_{2,k}^{1/5}| < \epsilon$ (ϵ small) to avoid the singularity.

5.3 Conclusion

In this Chapter, we have presented a solution to the $\mathcal{H}_2$ filtering problem for affine nonlinear descriptor systems in both continuous-time and discrete-time. Two types of filters have been presented; namely, singular and normal filters. Reduced-order normal filters have also been presented for the case of standard systems. Sufficient conditions for the solvability of the problem using each type of filter are given in terms of HJBEs and DHJBEs, and the results have also been specialized to linear systems, in which case the conditions reduce to a system of LMIs which are computationally efficient to solve. The problem for a nonconstant singular derivative matrix has also been discussed. Examples and simulation results have also been presented to illustrate the approach.

CHAPTER 6

$\mathcal{H}_\infty$ FILTERING FOR DESCRIPTOR NONLINEAR SYSTEMS

In this chapter, we discuss the solution to the $\mathcal{H}_\infty$ filtering problem for affine nonlinear descriptor systems. The corresponding $\mathcal{H}_2$ solution has been discussed in Chapter 5. This approach is useful when the disturbances/measurement noise in the system are known to be Gaussian. However, in the presence of nonGaussian noise and possibly $\mathcal{L}_2$-bounded disturbances, the $\mathcal{H}_\infty$ methods that we discuss in this chapter are more effective.

The chapter is organized as follows. In Section 2, we present a solution to the continuous-time problem. Two classes of filters, namely, (i) singular; and (ii) normal filters, are similarly considered. The general problem of a nonconstant derivative matrix is also considered. Examples are then presented to demonstrate the approach. Then in Section 3, we present the corresponding solution for discrete-time systems. Finally in Section 4, we give a short conclusion.

6.1 $\mathcal{H}_\infty$ Filtering for Continuous-time Systems

In this section, we present a solution to the continuous-time filtering problem using singular and normal filters. We begin with the problem definition and other preliminaries.

6.1.1 Problem Definition and Preliminaries

The general set-up for studying $\mathcal{H}_\infty$ filtering problems is shown in Fig. 6.1, where **P** is the plant, while **F** is the filter. The noise signal $w \in \mathcal{P}$ is in general a bounded power signal (or $\mathcal{L}_2$ signal) which belongs to the set $\mathcal{P}$ of bounded power signals, and similarly $\tilde{z} \in \mathcal{P}$, is a bounded power signal. Thus, the induced norm from w to $\tilde{z}$ (the penalty variable to be defined later) is the $\mathcal{H}_\infty$-norm (or $\mathcal{L}_2$-gain) of the interconnected system $\mathbf{F} \circ \mathbf{P}$, i.e.,

$$\|\mathbf{F} \circ \mathbf{P}\|_{\mathcal{H}_\infty} \stackrel{\Delta}{=} \sup_{0 \neq w \in \mathcal{L}_2} \frac{\|\tilde{z}\|_{\mathcal{L}_2}}{\|w\|_{\mathcal{L}_2}}. \tag{6.1}$$

At the outset, we consider the following affine nonlinear causal descriptor model of the plant, which is defined on a manifold $\mathcal{X} \subseteq \Re^n$ with zero control input:

$$\mathbf{P}_D^a : \begin{cases} E\dot{x} &= f(x) + g_1(x)w; \quad x(t_0) = x_0 \\ y &= h_2(x) + k_{21}(x)w, \end{cases} \tag{6.2}$$

where $x \in \mathcal{X}$ is the semistate vector; $w \in \mathcal{W} \subset \Re^r$ is an unknown disturbance (or noise) signal, which belongs to the set $\mathcal{W}$ of admissible exogenous inputs; $y \in \mathcal{Y} \subset \Re^m$ is the measured output (or observation) of the system, and belongs to $\mathcal{Y}$, the set of admissible measured-outputs.

The functions $f : \mathcal{X} \to T\mathcal{X}$ [1], $g_1 : \mathcal{X} \to \mathcal{M}^{n_1 \times r}(\mathcal{X})$, where $\mathcal{M}^{i \times j}$ is the ring of $i \times j$ smooth matrices over $\mathcal{X}$, $h_2 : \mathcal{X} \to \Re^m$, and $k_{21} : \mathcal{X} \to \mathcal{M}^{m \times r}(\mathcal{X})$ are real C^∞ functions of x, while $E \in \Re^{q \times n}$ is a constant but generally singular matrix. Furthermore, we assume without any loss of generality that the system (6.2) has an isolated equilibrium-point at $x = 0$ such that $f(0) = 0$, $h_2(0) = 0$. We also assume that there exists at least one solution $x(t, t_0, Ex_0, w)$ $\forall t \in \Re$ for the system for all admissible initial conditions Ex_0, for all $w \in \mathcal{W}$. The initial condition Ex_0 is said to be admissible if the solution $x(t)$ is unique, impulse-free and smooth for all $[t_0, \infty)$. In addition, the following standing assumptions will be made on the system.

For simplicity we also make the following assumptions on the plant.

Assumption 6.1.1. *The system matrices are such that*

$$k_{21}(x) g_1^T(x) = 0,$$
$$k_{21}(x) k_{21}^T(x) = I.$$

[1]For a manifold M, TM and $T^\star M$ are the tangent and cotangent bundles of M.

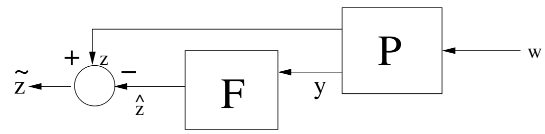

Figure 6.1 Set-up for $\mathcal{H}_\infty$ Filtering

Assumption 6.1.2. *Let $\bar{x} \in \mathcal{O} \subset \mathcal{X}$ a neighborhood of $x = 0$, $A = \frac{\partial f}{\partial x}(\bar{x})$. Then, the system (6.2) is locally admissible, implies the following hold:*

1. *the system is locally regular at each $\bar{x} \in \mathcal{O}$ and hence locally solvable, i.e, $det(sE-A) \not\equiv 0$ for all $s \in \mathbf{C}$;*
2. *the system is locally impulse-free at each $\bar{X} \in \mathcal{O}$, i.e., $deg(det(sE-A)) = rank(E)$ for all $\bar{x} \in \mathcal{O}$ and al most all $s \in \mathbf{C}$;*
3. *the system is locally asymptotically stable, i.e., (E, A) is Hurwitz for all $\bar{x} \in \mathcal{O}$.*

The suboptimal $\mathcal{H}_\infty$ local filtering/state-estimation problem is defined as follows.

Definition 6.1.1. *(Suboptimal $\mathcal{H}_\infty$ Local State Estimation or Filtering Problem). Find a filter, $\mathbf{F}$, for estimating the state $x(t)$ or a function of it, $z = h_1(x)$, from observations $\mathsf{Y}_t \triangleq \{y(\tau) : \tau \leq t\}$ of $y(\tau)$ up to time t, to obtain the estimate*

$$\hat{x}(t) = \mathbf{F}(\mathsf{Y}_t),$$

such that, the $\mathcal{H}_\infty$-norm (or $\mathcal{L}_2$-gain) from the input w to some suitable penalty function $\tilde{z}$ is locally less or equal to some given desired number $\gamma > 0$ for all initial conditions $Ex_0 \in \mathcal{O} \subset \mathcal{X}$. Moreover, if the filter solves the problem for all admissible $Ex_0 \in \mathcal{X}$, we say the problem is solved globally.

We shall adopt the following definition of local zero-input observability (Ozcaldiran, 1992).

Definition 6.1.2. *For the nonlinear system $\mathbf{P}^a_D$, we say that it is locally weakly zero-input* *observable, if for all states* $x_1,\ x_2 \in U \subset \mathcal{X}$ *and input* $w(.) = 0,\ t > t_0$

$$y(t; Ex_1(t_0-), w) \equiv y(t; Ex_2(t_0-), w) \Longrightarrow Ex_1(t_0) = Ex_2(t_0); \tag{6.3}$$

the system is said to be locally zero-input observable if

$$y(t; Ex_1(t_0-), w) \equiv y(t; Ex_2(t_0-), w) \Longrightarrow x_1(t_0) = x_2(t_0); \tag{6.4}$$

where $y(., Ex_i(t_0-), w),\ \ i = 1, 2$ *is the output of the system with the initial condition* $Ex_i(t_0-)$*; and the system is said to be locally strongly zero-input observable if*

$$y(t; Ex_1(t_0-), w) \equiv y(t; Ex_2(t_0-), w) \Longrightarrow x_1(t_0-) = x_2(t_0-). \tag{6.5}$$

Moreover, the system is said to be globally (weakly, strongly) zer-input observable, if it is locally (weakly, strongly) zero-input observable at each $x_0 \in \mathcal{X}$ *or* $U = \mathcal{X}$*.*

In the sequel, we shall not distinguish between observability and strong observability. Moreover, we shall also assume throughout that the noise signal $w \in \mathcal{W} \subset \mathcal{L}_2([t_0, \infty); \Re^r)$.

6.1.2 $\mathcal{H}_\infty$ Singular Filters

In this subsection, we discuss full-order $\mathcal{H}_\infty$ singular filters for the system in the usual Kalman-Luenberger type structure:

$$\mathbf{F}^a_{DS1} \begin{cases} E\dot{\hat{x}} &= f(\hat{x}) + g_1(\hat{x})\hat{w}^\star + \hat{L}(\hat{x}, y)(y - h_2(\hat{x}) - k_{21}(\hat{x})\hat{w}^\star) \\ \tilde{z} &= y - h_2(\hat{x}) \end{cases} \tag{6.6}$$

where $\hat{x} \in \mathcal{X}$ is the filter state, $\hat{w}^\star$ is the worst-case estimated system noise (or certainty-equivalent noise), $\hat{L} : \mathcal{X} \times \mathcal{Y} \to \Re^{n \times m}$ is the gain matrix of the filter, $\tilde{z} \in \Re^m$ is the penalty variable, or innovation variable, or estimation error. The problem can then be formulated as a dynamic optimization problem (Basar, 1995) with the following cost functional

$$\begin{aligned} \min_{\hat{L}\in\Re^{n\times m}} \sup_{w\in\mathcal{W}} \hat{J}(\hat{L},w) &= \frac{1}{2}\int_{t_0}^{\infty}(\|\tilde{z}(t)\|^2-\gamma^2\|w(t)\|)dt, \quad s.t.\ (6.6) \\ & \text{and with } w=0, \ \lim_{t\to\infty}\{\hat{x}(t)-x(t)\}=0. \end{aligned} \tag{6.7}$$

To solve the above problem, we form the Hamiltonian function $H : T^{\star}\mathcal{X}\times T^{\star}\mathcal{Y}\times\mathcal{W}\times\Re^{n\times m}\to\Re$:

$$\begin{aligned} H(\hat{x},y,w,\hat{L},\hat{V}_{E\hat{x}}^T,\hat{V}_y^T) &= \hat{V}_{E\hat{x}}(E\hat{x},y)[f(\hat{x})+g_1(\hat{x})w+\hat{L}(\hat{x},y)(y-h_2(\hat{x})-k_{21}(\hat{x})w)+ \\ & \quad \hat{V}_y(E\hat{x},y)\dot{y}+\frac{1}{2}(\|\tilde{z}\|^2-\gamma^2\|w\|^2) \end{aligned} \tag{6.8}$$

for some C^1 function $\hat{V}:\mathcal{X}\times\mathcal{Y}\to\Re$, and where $\hat{V}_{E\hat{x}}$ is the row vector of first partial-derivatives of $\hat{V}$ with respect to $E\hat{x}$. Applying then the necessary condition for the worst-case noise/disturbance, $\frac{\partial H}{\partial w}\big|_{w=\hat{w}^\star}=0$, we get

$$\hat{w}^\star=\frac{1}{\gamma^2}[g_1^T(\hat{x})-k_{21}^T(\hat{x})\hat{L}^T(\hat{x},y)]\hat{V}_{E\hat{x}}^T(E\hat{x},y),$$

and substituting $\hat{w}^\star$ in (6.8), we get

$$\begin{aligned} H(\hat{x},y,\hat{w}^\star,\hat{L},\hat{V}_{E\hat{x}}^T,\hat{V}_y^T) &= \hat{V}_{E\hat{x}}(E\hat{x},y)f(\hat{x})+\hat{V}_y(E\hat{x},y)\dot{y}+ \\ & \quad \frac{1}{2\gamma^2}\hat{V}_{E\hat{x}}(E\hat{x},y)g_1(\hat{x})g_1^T(\hat{x})\hat{V}_{E\hat{x}}^T(E\hat{x},y)+ \\ & \quad \hat{V}_{E\hat{x}}(E\hat{x},y)\hat{L}(\hat{x},y)(y-h_2(\hat{x}))+ \\ & \quad \frac{1}{2\gamma^2}\hat{V}_{E\hat{x}}(E\hat{x},y)\hat{L}(\hat{x},y)\hat{L}^T(\hat{x},y)\hat{V}_{E\hat{x}}^T(E\hat{x},y)+\frac{1}{2}\|\tilde{z}\|^2. \end{aligned}$$

Completing the squares now for $\hat{L}$ in the above expression (6.8), we have

$$\begin{aligned} H(\hat{x},y,\hat{w}^\star,\hat{L},\hat{V}_{E\hat{x}}^T,\hat{V}_y^T) &= \hat{V}_{E\hat{x}}(E\hat{x},y)f(\hat{x})+\hat{V}_y(E\hat{x},y)\dot{y}+ \\ & \quad \frac{1}{2\gamma^2}\hat{V}_{E\hat{x}}(E\hat{x},y)g_1(\hat{x})g_1^T(\hat{x})\hat{V}_{E\hat{x}}^T(E\hat{x},y) \\ & \quad +\frac{1}{2\gamma^2}\|\hat{L}^T(\hat{x},y)\hat{V}_{E\hat{x}}^T(E\hat{x},y)+\gamma^2(y-h_2(\hat{x}))\|^2+\frac{(1-\gamma^2)}{2}\|\tilde{z}\|^2. \end{aligned}$$

Thus, setting the optimal gain $\hat{L}^\star(\hat{x}, y)$ as

$$\hat{V}_{E\hat{x}}(E\hat{x}, y)\hat{L}^\star(\hat{x}, y) = -\gamma^2(y - h_2(\hat{x}))^T, \tag{6.9}$$

minimizes the Hamiltonian (6.8) and ensures that the saddle-point condition (Basar, 1995)

$$H(.,.,\hat{w}^\star, \hat{L}^\star, .,.) \leq H(.,.,\hat{w}^\star, \hat{L}, .,.) \tag{6.10}$$

is satisfied.

In addition, notice similarly as in Chapter 5, from (6.2) and with the measurement noise set at zero,

$$\dot{y} = \tilde{\mathcal{L}}_{f+gw} h_2,$$

where $\tilde{\mathcal{L}}$ is the Lie-derivative operator (Sastry, 1999) in coordinates Ex. Moreover, under certainty-equivalence and using the expression for $\hat{w}^\star$ as defined above, we have

$$\dot{y} = \tilde{\mathcal{L}}_{f(\hat{x})+g_1(\hat{x})\hat{w}^\star} h_2(\hat{x}) = \nabla_{E\hat{x}} h_2(\hat{x})[f(\hat{x}) + \frac{1}{\gamma^2} g_1(\hat{x}) g_1^T(\hat{x}) \hat{V}_{E\hat{x}}^T(E\hat{x}, y)].$$

Finally, setting

$$H(\hat{x}, y, \hat{w}^\star, \hat{L}^\star, \hat{V}_{E\hat{x}}^T, \hat{V}_y^T) = 0$$

results in the following Hamilton-Jacobi-Isaac's equation (HJIE):

$$\begin{gathered}
\hat{V}_{E\hat{x}}(E\hat{x}, y) f(\hat{x}) + \hat{V}_y(E\hat{x}, y)\nabla_{E\hat{x}} h_2(\hat{x}) f(\hat{x}) + \\
\tfrac{1}{\gamma^2}\hat{V}_y(E\hat{x}, y)\nabla_{E\hat{x}} h_2(\hat{x}) g_1(\hat{x}) g_1^T(\hat{x}) \hat{V}_{E\hat{x}}^T(E\hat{x}, y) + \\
\tfrac{1}{2\gamma^2}\hat{V}_{E\hat{x}}(E\hat{x}, y) g_1(\hat{x}) g_1^T(\hat{x}) \hat{V}_{E\hat{x}}^T(E\hat{x}, y) + \\
\tfrac{(1-\gamma^2)}{2}(y - h_2(\hat{x}))^T (y - h_2(\hat{x})) = 0, \quad \hat{V}(0,0) = 0.
\end{gathered} \tag{6.11}$$

We then have the following result.

Proposition 6.1.1. *Consider the nonlinear system (6.2) and the local $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\mathbf{P}_\mathbf{D}^\mathbf{a}$ satisfies Assumption 6.1.1, is locally asymptotically* *stable about the equilibrium-point $x = 0$, and zero-input observable. Further, suppose for some $\gamma > 0$, there exist a C^1 positive-semidefinite function $\hat{V} : \hat{N} \times \hat{\Upsilon} \to \Re_+$ locally defined in a neighborhood $\hat{N} \times \hat{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\hat{x}, y) = (0, 0)$, and a matrix function $\hat{L} : \hat{N} \times \hat{\Upsilon} \to \Re^{n\times m}$, satisfying the HJIE (6.11) together with the side-condition (6.9). Then the filter $\mathbf{F}^a_{DS1}$ solves the local $\mathcal{H}_\infty$ filtering problem for the system.*

Proof: To complete the proof, we need to show that $(\hat{w}^\star, L^\star)$ constitute a saddle-point solution to the optimization problem (6.7), i.e,

$$H(.,., w, \hat{L}^\star, ., .) \leq H(.,., \hat{w}^\star, \hat{L}^\star, ., .) \leq H(.,., \hat{w}^\star, \hat{L}, ., .) \tag{6.12}$$

hold, and the $\mathcal{L}_2$-gain condition

$$\|\tilde{z}\|_{\mathcal{L}_2} \leq \gamma^2 \|w\|_{\mathcal{L}_2} \tag{6.13}$$

is satisfied with $L = \hat{L}^\star$.

First, note from (6.8) and (6.11)

$$\begin{aligned} H(\hat{x}, y, w, \hat{L}^\star, \hat{V}^T_{\hat{x}}, \hat{V}^T_y) &= \Big\{ \hat{V}_{E\hat{x}}(E\hat{x}, y) f(\hat{x}) + \hat{V}_y(E\hat{x}, y)\dot{y} + \\ &\quad \frac{1}{2\gamma^2}\hat{V}_{E\hat{x}}(E\hat{x}, y) g_1(\hat{x}) g_1^T(\hat{x}) \hat{V}^T_{E\hat{x}}(E\hat{x}, y) + \frac{(1-\gamma^2)}{2}\|\tilde{z}\|^2 \Big\} - \\ &\quad \frac{\gamma^2}{2}\|w - \hat{w}^\star\|^2 \\ &= H(\hat{x}, y, \hat{w}^\star, \hat{L}^\star, \hat{V}^T_{\hat{x}}, \hat{V}^T_y) - \frac{\gamma^2}{2}\|w - \hat{w}^\star\|^2. \end{aligned}$$

Therefore,

$$H(\hat{x}, y, \hat{w}, \hat{L}^\star, \hat{V}^T_{\hat{x}}, \hat{V}^T_y) \leq H(\hat{x}, y, \hat{w}^\star, \hat{L}^\star, \hat{V}^T_{\hat{x}}, \hat{V}^T_y). \tag{6.14}$$

Hence, combining (6.10) and (6.14), we have that the sadle-point conditions (6.12) are satisfied.

Next, let $\hat{V} \geq 0$ be a C^1 solution of the HJIE (6.11). Then, differentiating this solution along a trajectory of (6.6), with $\hat{L} = \hat{L}^\star$, we get

$$\begin{aligned}
\dot{\hat{V}} &= \hat{V}_{E\hat{x}}(E\hat{x}, y)[f(\hat{x}) + g_1(\hat{x})w + \hat{L}^\star(\hat{x}, y)(y - \tilde{h}_2(\hat{x}))] + \hat{V}_y(E\hat{x}, y)\dot{y} \\
&= \Big\{\hat{V}_{E\hat{x}}(E\hat{x}, y)f(\hat{x}) + \hat{V}_y(E\hat{x}, y)\dot{y} + \frac{1}{2\gamma^2}\hat{V}_{E\hat{x}}(E\hat{x}, y)g_1(\hat{x})g_1^T(\hat{x})\hat{V}_{E\hat{x}}^T(E\hat{x}, y) + \\
&\quad \frac{(1-\gamma^2)}{2}\|\tilde{z}\|^2\Big\} - \frac{\gamma^2}{2}\|w - \hat{w}^\star\|^2 + \frac{\gamma^2}{2}\|w\|^2 - \frac{1}{2}\|\tilde{z}\|^2 \\
&\leq \frac{\gamma^2}{2}\|w\|^2 - \frac{1}{2}\|\tilde{z}\|^2,
\end{aligned}$$

where the last inequality follows from using the HJIE (6.11). Moreover, setting $w = 0$ in the above inequality we get $\dot{\hat{V}} \leq -\frac{1}{2}\|\tilde{z}\|^2$. Therefore, the filter dynamics is stable, and $\hat{V}(E\hat{x}, y)$ is non-increasing along a trajectory of (6.6). Further, the condition that $\dot{\hat{V}}(E\hat{x}(t), y(t)) \equiv 0\ \forall t \geq t_s$ implies that $\tilde{z} \equiv 0$, which further implies that $y = h_2(\hat{x})\ \forall t \geq t_s$. By the zero-input observability of the system, this implies that $\hat{x} = x\ \forall t \geq t_s$.

Finally, integrating the above inequality (6.15) from $t = t_0$ to $t = \infty$, and since $\hat{V}(E\hat{x}(\infty), y(\infty)) < \infty$, we get that the $\mathcal{L}_2$-gain condition (6.13) is satisfied. □

The result of the theorem can be specialized to the linear descriptor system

$$\mathbf{P}_D^l : \begin{cases} E\dot{x} &= Ax + B_1 w; \quad Ex(t_0) = Ex_0 \\ y &= C_2 x + D_{21} w, \end{cases} \tag{6.15}$$

where $E \in \Re^{n\times n}$, $A \in \Re^{n\times n}$, $B_1 \in \Re^{n\times r}$, $C_2 \in \Re^{m\times n}$, $D_{21} \in \Re^{m\times r}$.

Corollary 6.1.1. *Consider the linear descriptor system (6.15) and the $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\mathbf{P}_D^l$ is locally asymptotically stable about the equilibrium-point $x = 0$ and observable. Suppose further, for some $\gamma > 0$, there exist symmetric positive-semidefinite matrices $\hat{P} \in \Re^{n\times n}$, $\hat{Q} \in \Re^{m\times m}$, and a matrix $\hat{L} \in \Re^{n\times m}$, satisfying the LMIs*

$$\begin{bmatrix} E^T\hat{P}A + A^T\hat{P}E + (1-\gamma^2)C_2^T C_2 & E^T\hat{P}B_1 & (1-\gamma^2)C_2^T & 0 \\ B_1^T\hat{P}E & -\gamma^2 I & 0 & 0 \\ (1-\gamma^2)C_2 & 0 & -I & \hat{Q} \\ 0 & 0 & \hat{Q}^T & 0 \end{bmatrix} \leq 0 \tag{6.16}$$

$$\begin{bmatrix} 0 & \frac{1}{2}(E^T\hat{P}\hat{L}-\gamma^2 C_2^T) \\ \frac{1}{2}(E^T\hat{P}\hat{L}-\gamma^2 C_2^T)^T & (1-\delta_1)I \end{bmatrix} \leq 0 \qquad (6.17)$$

for some number $\delta_1 > 1$. Then the filter

$$\mathbf{F}^l_{DS1}\ :\ \begin{cases} E\dot{\hat{x}} &= (A+\frac{1}{\gamma^2}B_1B_1^T\hat{P}E+\frac{1}{\gamma^2}\hat{L}\hat{L}^T\hat{P}E)\hat{x}+\hat{L}(y-C_2\hat{x}); \quad E\hat{x}(t_0)=0 \\ \hat{z} &= C_2\hat{x} \end{cases} \qquad (6.18)$$

solves the $\mathcal{H}_\infty$ estimation problem for the system.

Proof: Take $\hat{V}=\frac{1}{2}(\hat{x}^TE^T\hat{P}Ex+y^T\hat{Q}y)$ and apply the result of the Proposition. □

Furthermore, to improve the steady-state estimation error, we propose in addition a proportional-integral (PI) filter configuaration (Gao, 2004), (Koenig, 1995):

$$\mathbf{F}^a_{DS2}\begin{cases} E\dot{\check{x}} &= f(\check{x})+g_1(\check{x})\check{w}^\star+\check{L}_1(E\check{x},\xi,y)(y-h_2(\check{x})-k_{21}(\check{x})\check{w}^\star)+ \\ & \quad \check{L}_2(E\check{x},\xi,y)\xi, \quad E\check{x}(t_0)=0 \\ \dot{\xi} &= y-h_2(\check{x}) \\ \check{z} &= y-h_2(\check{x}) \end{cases} \qquad (6.19)$$

where $\check{x}\in\mathcal{X}$ is the filter state, $\xi\in\Re^m\times\Re$ is the integrator state, and $\check{L}_1,\check{L}_2:\mathcal{X}\times\mathcal{Y}\to\Re^{n\times m}$ are the proportional and integral gain matrices of the filter respectively. Similarly, using manipulations as in Proposition 6.1.1, we can arrive at the following result.

Theorem 6.1.1. *Consider the nonlinear system (6.2) and the local $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\mathbf{P}^{\mathbf{a}}_{\mathbf{D}}$ is locally asymptotically stable about the equilibrium-point $x=0$ and zero-input observable. Further, suppose for some $\gamma>0$, there exist a C^1 positive-semidefinite function $\check{V}:\check{N}\times\check{\Xi}\times\check{\Upsilon}\to\Re_+$ locally defined in a neighborhood $\check{N}\times\check{\Xi}\times\hat{\Upsilon}\subset\mathcal{X}\times\mathcal{Y}\times\Re\times\mathcal{Y}$ of the origin $(\check{x},\xi,y)=(0,0,0)$, and matrix functions $\check{L}_1,\check{L}_2:\check{N}\times\check{\Xi}\times\check{\Upsilon}\to\Re^{n\times m}$, satisfying the HJIE*

$$\check{V}_{E\check{x}}(E\check{x},\xi,y)f(\check{x})+\check{V}_y(E\check{x},\xi,y)\nabla_{E\check{x}}h_2(\check{x})f(\check{x})+$$
$$\frac{1}{\gamma^2}\check{V}_y(E\check{x},\xi,y)\nabla_{E\check{x}}h_2(\check{x})g_1(\check{x})g_1^T(\check{x})\check{V}^T_{E\check{x}}(E\check{x},y))+$$

$$
\begin{aligned}
\tfrac{1}{2\gamma^2}\check{V}_{E\check{x}}(E\check{x},\xi,y)g_1(\check{x})g_1^T(\check{x})\check{V}_{E\check{x}}^T(E\check{x},\xi,y)+ \\
\check{V}_\xi(E\check{x},\xi,y)(y-h_2(\hat{x}))-\xi^T\xi+ \\
\tfrac{(1-\gamma^2)}{2}(y-h_2(\check{x}))^T(y-h_2(\check{x}))=0,\quad \check{V}(0,0,0)=0,
\end{aligned}
\tag{6.20}
$$

together with the side-conditions

$$
\begin{aligned}
\check{V}_{E\check{x}}(E\check{x},\xi,y)\check{L}_1(E\check{x},\xi,y) &= -\gamma^2(y-h_2(\check{x}))^T, &(6.21)\\
\check{V}_{E\check{x}}(E\check{x},\xi,y)\check{L}_2(E\check{x},\xi,y) &= -\xi^T. &(6.22)
\end{aligned}
$$

Then, the filter $\mathbf{F}^a_{DS2}$ solves the $\mathcal{H}_\infty$ local filtering problem for the system.

In the next section, we consider the design of normal filters for the system.

6.1.3 $\mathcal{H}_\infty$ Normal Filters

In this subsection, we discuss normal filters for the system (6.2). We shall consider the design of both full-order and reduced-order filters. We start with the full-order filter first, and in this regard, without any loss of generality, we can assume that E is of the form

$$
E=\begin{pmatrix} I_{q\times q} & 0 \\ 0 & 0 \end{pmatrix}.
$$

This follows from matrix theory and can easily be proven using the singular-value decomposition (SVD) of E. It follows that, the system can be represented in the canonical form of a differential-algebraic system

$$
\bar{\mathbf{P}}^a_D : \begin{cases} \dot{x}_1 = f_1(x)+g_{11}(x)w; \;\; x(t_0)=x_0 \\ 0 = f_2(x)+g_{21}(x)w \\ y = h_2(x)+k_{21}(x)w, \end{cases} \tag{6.23}
$$

where $dim(x_1) = q$, $f_1(0) = 0$, $f_2(0) = 0$. Then, if we define

$$\dot{x}_2 = f_2(x) + g_{21}(x)w,$$

where $\dot{x}_2$ is a fictitious state vector, and $dim(x_2) = n-q$, the system (6.23) can be represented by a normal state-space system as

$$\widetilde{\mathbf{P}}^a_D : \begin{cases} \dot{x}_1 &= f_1(x) + g_{11}(x)w; \quad x_1(t_0) = x_{10} \\ \dot{x}_2 &= f_2(x) + g_{21}(x)w; \quad x_2(t_0) = x_{20} \\ y &= h_2(x) + k_{21}(x)w. \end{cases} \tag{6.24}$$

Now define the set

$$\Omega_o = \{(x_1, x_2) \in \mathcal{X} \mid \dot{x}_2 = 0\}. \tag{6.25}$$

Then, we have the following system equivalence

$$\widetilde{\mathbf{P}}^a_D|_{\Omega_o} = \bar{\mathbf{P}}^a_D. \tag{6.26}$$

Therefore, to estimate the states of the system (6.23), we need to stabilize the system (6.24) about Ω_o and then design a filter for the resulting system. For this purpose, we consider the following class of filters

$$\mathbf{F}^a_{DN3} \begin{cases} \dot{\grave{x}}_1 &= f_1(\grave{x}) + g_{11}(\grave{x})\grave{w}^\star + \grave{L}_1(\grave{x}, y)[y - h_2(\grave{x}) - k_{21}(\grave{x})\grave{w}^\star] \\ \dot{\grave{x}}_2 &= f_2(\grave{x}) + g_{21}(\grave{x})\grave{w}^\star + g_{22}(x)\alpha_2(\grave{x}) + \grave{L}_2(\grave{x}, y)[y - h_2(\grave{x}) - k_{21}(\grave{x})\grave{w}^\star] \\ \grave{z} &= y - h_2(\grave{x}), \end{cases} \tag{6.27}$$

where $\grave{x} \in \mathcal{X}$ is the filter state, $\grave{w}^\star$ is the estimated worst-case system noise, $\grave{L}_1 : \mathcal{X} \times \mathcal{Y} \to \Re^{q\times m}$, $\grave{L}_2 : \mathcal{X} \times \mathcal{Y} \to \Re^{n-q\times m}$ are the filter gain matrices, and $g_{22} : \mathcal{X} \to \mathcal{M}^{(n-q)\times p}$ is a gain matrix for the artificial control input $u = \alpha_2(x) \in \Re^p$ required to stabilize the dynamics $\dot{x}_2$ about Ω_o. Accordingly, we make the following assumption.

Assumption 6.1.3. *The pair $\{f_2, g_{22}\}$ is stabilizable, i.e., $\exists$ a control-Lyapunov-function (CLF), $\bar{V} > 0$, such that $\bar{V}_{x_2}(x)(f_2(x) - g_{22}(x)g_{22}^T(x)\bar{V}_{x_2}^T(x)) < 0$.*

Similarly, for simplicity in the derivation of the results, we also make the following assumption on the plant $\bar{\mathbf{P}}^a_D$.

Assumption 6.1.4. *The system matrices are such that*

$$k_{21}(x)[g_{11}^T(x) \;\; g_{21}^T(x)] = 0,$$
$$k_{21}(x)k_{21}^T(x) = I.$$

Thus, if Assumption 6.1.3 holds, then we can set $\alpha_2(\grave{x}) = -\frac{1}{\varepsilon}g_{22}^T(\grave{x})\bar{V}_{\grave{x}_2}^T(\grave{x})$, where $\varepsilon > 0$ is small, a high-gain feedback (Young, 1977) to constrain the dynamics on Ω_o as fast as possible. Then we proceed to design the gain matrices $\grave{L}_1$, $\grave{L}_2$ to estimate the states. Moreover, a common HJI-CLF can also be utilized in the above design procedure for both the stabilization and the filtering. Consequently, we have the following result.

Proposition 6.1.2. *Consider the nonlinear system (6.23) and the local $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\bar{\mathbf{P}}^{\mathbf{a}}_{\mathbf{D}}$ satisfies Assumptions 6.1.3, 6.1.4, is locally asymptotically stable about the equilibrium-point $x = 0$, and zero-input observable. Further, suppose for some $\gamma > 0$, there exist a C^1 positive-semidefinite function $\grave{V} : \grave{N} \times \grave{\Upsilon} \to \Re_+$, locally defined in a neighborhood $\grave{N} \times \grave{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\grave{x}_1, \grave{x}_2, y) = (0, 0, 0)$, and matrix functions $\grave{L}_1 : \grave{N} \times \grave{\Upsilon} \to \Re^{q\times m}$, $\grave{L}_2 : \grave{N} \times \grave{\Upsilon} \to \Re^{n-q\times m}$, satisfying the HJIE:*

$$\begin{aligned}
&\grave{V}_{\grave{x}_1}(\grave{x},y)f_1(\grave{x}) + \grave{V}_{\grave{x}_2}(\grave{x},y)f_2(\grave{x}) + \grave{V}_y(\grave{x},y)\nabla_{\grave{x}_1}h_2(\grave{x})f_1(\grave{x}) + \\
&\tfrac{1}{\gamma^2}\grave{V}_y(\grave{x},y)\nabla_{\grave{x}_1}h_2(\grave{x})[g_{11}(\grave{x})g_{11}^T(\grave{x})\grave{V}_{\grave{x}_1}^T(\grave{x},y) + g_{11}(\grave{x})g_{21}^T(\grave{x})\grave{V}_{\grave{x}_2}^T(\grave{x},y)] + \\
&\tfrac{1}{2\gamma^2}\grave{V}_{\grave{x}_1}(\grave{x},y)g_{11}(\grave{x})g_{11}^T(\grave{x})\grave{V}_{\grave{x}_1}^T(\grave{x},y) + \tfrac{1}{\gamma^2}\grave{V}_{\grave{x}_1}(\grave{x},y)g_{11}(\grave{x})g_{21}^T(\grave{x})\grave{V}_{\grave{x}_2}^T(\grave{x},y) + \\
&\tfrac{1}{2\gamma^2}\grave{V}_{\grave{x}_2}(\grave{x},y)g_{21}(\grave{x})g_{21}^T(\grave{x})\bar{V}_{\grave{x}_2}^T(\grave{x},y) - \tfrac{1}{\varepsilon}\grave{V}_{\grave{x}_2}(\grave{x},y)g_{22}(\grave{x})g_{22}^T(\grave{x})\grave{V}_{\grave{x}_2}^T(\grave{x},y) + \\
&\tfrac{(1-\gamma^2)}{2}(y - h_2(\grave{x}))^T(y - h_2(\grave{x})) = 0, \quad \grave{V}(0,0) = 0,
\end{aligned} \tag{6.28}$$

together with the side-conditions

$$\grave{V}_{\grave{x}_1}(\grave{x},y)\grave{L}_1(\grave{x},y) + \grave{V}_{\grave{x}_2}(\grave{x},y)\grave{L}_2(\grave{x},y) = -\gamma^2(y - h_2(\grave{x}))^T. \tag{6.29}$$

Then, the filter $\mathbf{F}^a_{DN3}$ solves the local $\mathcal{H}_\infty$-filtering problem for the system.

Proof: Follows along same lines as Proposition 6.1.1.

Remark 6.1.1. *Notice the addition of the high-gain feedback $u = \alpha_2(\check{x})$, transforms the filter $\mathbf{F}^a_{DN3}$ to a singularly-perturbed system (Young, 1977) with a slow subsystem governed by the dynamics $\dot{x}_1$, and a fast subsystem governed by the x_2-dynamics. This also suggests an alternative approach to the filter design problem, by considering a singularly-perturbed model of the system (6.23) as*

$$\tilde{\mathbf{P}}^a_{\varepsilon D} : \begin{cases} \dot{x}_1 &= f_1(x) + g_{11}(x)w; \quad x(t_0) = x_0 \\ \varepsilon\dot{x}_2 &= f_2(x) + g_{21}(x)w, \\ y &= h_2(x) + k_{21}(x)w, \end{cases} \tag{6.30}$$

where $\varepsilon > 0$ is a small parameter, and designing a normal filter for this equivalent system (Aliyu, 2011b). Notice in this case, as $\varepsilon \to 0$, the model (6.30) reduces to the original model (6.23).

Similarly, a normal PI-filter for the system (6.23) can also be designed. However, next we consider a reduced-order normal filter design. Accordingly, partition the state-vector x comformably with $rank(E) = q$ as $x = (x_1^T \;\; x_2^T)^T$ with $dim(x_1) = q$, $dim(x_2) = n - q$ and the state equations as

$$\breve{\mathbf{P}}^a_{D} : \begin{cases} \dot{x}_1 &= f_1(x_1, x_2) + g_{11}(x_1, x_2)w; \quad x_1(t_0) = x_{10} \\ 0 &= f_2(x_1, x_2) + g_{21}(x_1, x_2)w; \quad x_2(t_0) = x_{20} \\ y &= h_2(x) + k_{21}(x)w. \end{cases} \tag{6.31}$$

Then we make the following assumption.

Assumption 6.1.5. *The system is in the standard-form, i.e., the Jacobian matrix $f_{2,x_2}(x_1, x_2)$ is nonsigular in an open neighborhood $\widetilde{U}$ of $(0, 0)$ and $g_{21}(0, 0) \neq 0$.*

If Assumption 6.1.5 holds, then by the Implicit-function Theorem (Sastry, 1999), there exists a unique C^1 function $\phi : \Re^q \times \mathcal{W} \to \Re^{n-q}$ and a solution

$$\bar{x}_2 = \phi(x_1, w)$$

to equation (6.31b). Therefore, the system can be locally reprsented in $\widetilde{U}$ as the reduced-order system

$$\bar{\mathbf{P}}^a_{rD} : \begin{cases} \dot{x}_1 = f_1(x_1, \phi(x_1, w)) + g_{11}(x_1, \phi(x_1, w))w; \quad x_1(t_0) = x_{10} \\ y = h_2(x_1, \phi(x_1, w)) + k_{21}(x_1, \phi(x_1, w))w. \end{cases} \tag{6.32}$$

We can then design a normal filter of the form

$$\mathbf{F}^a_{DrN4} \begin{cases} \dot{\breve{x}}_1 = f_1(\breve{x}_1, \phi(\breve{x}_1, \breve{w}^\star)) + g_{11}(\breve{x}_1, \breve{x}_2)\breve{w}^\star + \breve{L}(\breve{x}_1, \phi(\breve{x}_1, \breve{w}^\star), y)[y- \\ \qquad h_2(\breve{x}_1, \phi(\breve{x}_1, \breve{w}^\star)) - k_{21}(\breve{x}_1, \phi(\breve{x}_1, \breve{w}^\star))]; \quad \breve{x}_1(t_0) = 0 \\ \breve{z} = y - h_2(\breve{x}_1, \phi(\breve{x}_1, \breve{w}^\star)) \end{cases} \tag{6.33}$$

for the system, and consequently, we have the following result.

Theorem 6.1.2. *Consider the nonlinear system (6.23) and the $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\bar{\mathbf{P}}^{\mathbf{a}}_{\mathbf{D}}$ satisfies Assumptions 6.1.3, 6.1.5, is locally asymptotically stable about the equilibrium-point $x = 0$, and zero-input observable. Further, suppose for some $\gamma > 0$, there exists a C^1 positive-semidefinite function $\breve{V} : \breve{N} \times \breve{\Upsilon} \to \Re_+$, locally defined in a neighborhood $\breve{N} \times \breve{\Upsilon} \subset \widetilde{U} \times \mathcal{Y}$ of the origin $(\breve{x}_1, y) = (0, 0)$, and a matrix function $\breve{L} : \breve{N} \times \breve{\Upsilon} \to \Re^{q\times m}$, satisfying the HJIE:*

$$\begin{aligned} &\breve{V}_{\breve{x}_1}(\breve{x}_1, y) f_1(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star)) + \grave{V}_y(\grave{x}, y)\nabla_{\grave{x}_1} h_2(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star)) f_1(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star)) + \\ &\frac{1}{2\gamma^2}\breve{V}_{\breve{x}_1}(\breve{x}_1, y) g_{11}(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star)) g_{11}^T(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star)\breve{V}_{\breve{x}_1}^T(\breve{x}_1, y) + \\ &\frac{1}{\gamma^2}\breve{V}_y(\grave{x}, y)\nabla_{\grave{x}_1} h_2(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star)) g_{11}(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star))) g_{11}^T(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star))\breve{V}_{\breve{x}_1}^T(\breve{x}_1, y) + \\ &\frac{(1-\gamma^2)}{2}(y - h_2(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star)))^T(y - h_2(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star))) = 0, \quad \breve{V}(0, 0) = 0, \end{aligned} \tag{6.34}$$

together with the side-conditions

$$\breve{w}^\star = \frac{1}{\gamma^2}[g_{11}^T(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star)) - k_{21}^T(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star))\breve{L}^T(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star)]\breve{V}_{\breve{x}_1}^T(\breve{x}_1, y),$$

$$\breve{V}_{\breve{x}_1}(\breve{x}_1, y)\breve{L}(\breve{x}_1, y) = -\gamma^2(y - h_2(\breve{x}_1, \varphi(\breve{x}_1, \breve{w}^\star))^T. \tag{6.35}$$

Then, the filter $\mathbf{F}^a_{DrN4}$ solves the local $\mathcal{H}_\infty$-filtering problem for the system.

Proof: Follows along same lines as Proposition 6.1.1.

Remark 6.1.2. *Notice in the above Theorem 6.1.2, $\breve{w}^\star$ is given implicitly, and so is the HJIE (6.34) given in terms of $\breve{w}^\star$. Thus, we can only find an approximate solution to the filtering problem.*

Similarly, we can specialize the result of Theorem 6.1.2 to the linear system (6.15). The system can be rewritten in the form (6.23) as

$$\mathbf{P}^l_D \ : \ \begin{cases} \dot{x} &= A_1x_1 + A_{12}x_2 + B_{11}w; \quad x_1(t_0) = x_{10} \\ 0 &= A_{21}x_1 + A_2x_2 + B_{21}w; \quad x_2(t_0) = x_{20} \\ y &= C_{21}x_1 + C_{22}x_2 + D_{21}w \end{cases} \tag{6.36}$$

Then, if A_2 is nonsingular, (Assumption 6.1.5) we can solve for x_2 in equation (6.36(b)) to get

$$\bar{x}_2 = -A_2^{-1}(A_{21}x_1 + B_{21}w),$$

and the filter (6.33) takes the following form

$$\mathbf{F}^l_{DrN4} \begin{cases} \dot{\breve{x}}_1 &= (A_1 - A_{12}A_2^{-1}A_{21})\breve{x}_1 + (B_{11} - A_{12}A_2^{-1}B_{21})\breve{w}^\star + \\ & \quad \breve{L}[y - (C_{21} - C_{22}A_2^{-1}A_{21})\breve{x}_1 - D_{21}\breve{w}^\star]; \quad \breve{x}_1(t_0) = 0 \\ \breve{z} &= y - (C_{21} - C_{22}A_2^{-1}A_{21})\breve{x}_1. \end{cases} \tag{6.37}$$

Then, we have the following corollary.

Corollary 6.1.2. *Consider the linear descriptor system (6.15) and the $\mathcal{H}_\infty$-filtering problem for this system. Suppose the plant $\mathbf{P}^l_D$ is locally asymptotically stable about the equilibrium-point $x = 0$, Assumption 6.1.5 holds and the plant is observable. Further, suppose for some $\gamma > 0$, there exist symmetric positive-semidefinite matrices $\breve{P} \in \Re^{q\times q}$, $\breve{Q} \in \Re^{m\times m}$, and a* *matrix* $\breve{L} \in \Re^{n\times m}$*, satisfying the LMIs:*

$$\begin{bmatrix} \breve{A}_1^T\breve{P} + \breve{P}\breve{A}_1 + (1-\gamma^2)\breve{C}_2^T\breve{C}_2^T & \breve{P}\breve{B}_1 & \gamma^2\breve{C}_2^T & 0 \\ \breve{B}_1^T\breve{P} & -\gamma^2 I & 0 & 0 \\ \gamma^2\breve{C}_2 & 0 & -I & \breve{Q} \\ 0 & 0 & \breve{Q} & 0 \end{bmatrix} \leq 0 \tag{6.38}$$

$$\begin{bmatrix} 0 & \frac{1}{2}(\breve{P}\breve{L} - \gamma^2\breve{C}_2^T) \\ \frac{1}{2}(\breve{P}\breve{L} - \gamma^2\breve{C}_2^T)^T & (1-\delta_3)I \end{bmatrix} \leq 0 \tag{6.39}$$

for some $\delta_3 \geq 1$*, where* $\breve{A}_1 = (A_1 - A_{12}A_2^{-1}A_{21})$*,* $\breve{B}_1 = (B_{11} - A_{12}A_2^{-1}B_{21})$*,* $\breve{C}_2 = (C_{21} - C_{22}A_2^{-1}A_{21})$*, and* $\breve{w}^\star = \frac{1}{\gamma^2}(\breve{B}_1^T - D_{12}^T\breve{L}^T)\breve{P}$*. Then the filter (6.37) solves the* $\mathcal{H}_\infty$*-filtering problem for the system.*

Proof Take$\breve{V}(\breve{x}) = \frac{1}{2}(\breve{x}_1^T\breve{P}\breve{x}_1 + y^T\breve{Q}y)$ and apply the result of the Theorem. □

In the next section, we consider a simple example due to the difficulty of solving the HJIE.

6.2 Examples

In this section, we consider a few simple examples due to space limitation.

Example 6.2.1. *Consider the following example of a nonlinear voltage controlled capacitor (Example III-2, (Newcomb, 1981b)) with* $C = 1F$*:*

$$\begin{aligned} \dot{x}_1 &= -x_3 \\ 0 &= -x_1 - x_2 + w \\ 0 &= -x_1 + 3x_3 + x_3^3 \\ y &= x_2. \end{aligned}$$

Eliminating the second equation, the above system can be represented as

$$\begin{aligned}\dot{x}_1 &= -x_3 \\ 0 &= -x_1 + 3x_3 + x_3^3 \\ y &= -x_1 + w.\end{aligned}$$

It can then be checked that for $\gamma = 1$*, the function* $\hat{V}(\hat{x}) = \frac{1}{2}\hat{x}_1^2$ *solves the inequality form of the HJIE (6.11), where the right-hand-side the HJIE reduces to* $\hat{V}_{E\hat{x}}(E\hat{x})f(\hat{x}) = -\hat{x}_1\hat{x}_3 = -3\hat{x}_3^2 - \hat{x}_3^4 \leq 0$*. Hence we can calculate the gain of the singular filter* $\mathbf{F}^a_{DS1}$ *as* $\hat{l}_1 = -\frac{(y+\hat{x}_1)}{\hat{x}_1}$.

Example 6.2.2. *Consider now a modified version of the voltage controlled capacitor of Example 6.2.1:*

$$\dot{x}_1 = -x_1^3 + x_2 \tag{6.40}$$

$$0 = -x_1 - x_2 \tag{6.41}$$

$$y = 2x_1 + x_2 + w \tag{6.42}$$

We find the gain for the singular filter $\mathbf{F}^a_{DS1}$ *presented in subsection 6.1.2. It can be checked that the system is locally observable, and the function* $\hat{V}(\hat{x}) = \frac{1}{2}\hat{x}_1^2$*, solves the inequality form of the HJIE (6.11) for the system. Then, we calculate the gain of the filter as*

$$\hat{L}(\hat{x}, y) = -\frac{(y - 2\hat{x}_1 - \hat{x}_2)}{\hat{x}_1},$$

where $\hat{L}$ *is set equal zero if* $|\hat{x}_1| < \epsilon$ *(*ϵ *small) to avoid the singularity at* $\hat{x}_1 = 0$.

Figures 6.2 and 6.3 show the result of the simulation with the above filter. In Figure 6.2, the noise is a uniformly distributed noise with variance of 0.2*, while in Figure 6.3 we have* $w(t) = e^{-0.2t}\sin(0.5\pi t)$ *is an* $\mathcal{L}_2$*-bounded disturbance. The result of the simulations show good convergence with unknown system initial conditions.*

Similarly, we can determine the reduced-order filter gain (6.35) for the above system. Notice that the system also satisfies Assumption 6.1.5, thus we can solve equation (6.41) for x_2 *to*

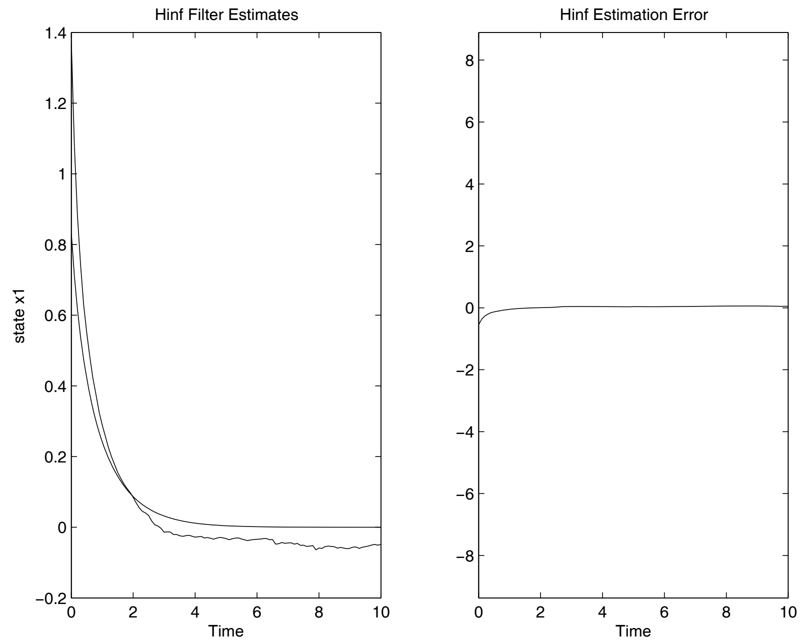

Figure 6.2 $\mathcal{H}_\infty$ *singular filter performance with unknown initial condition and uniformly distributed noise with variance* 0.2

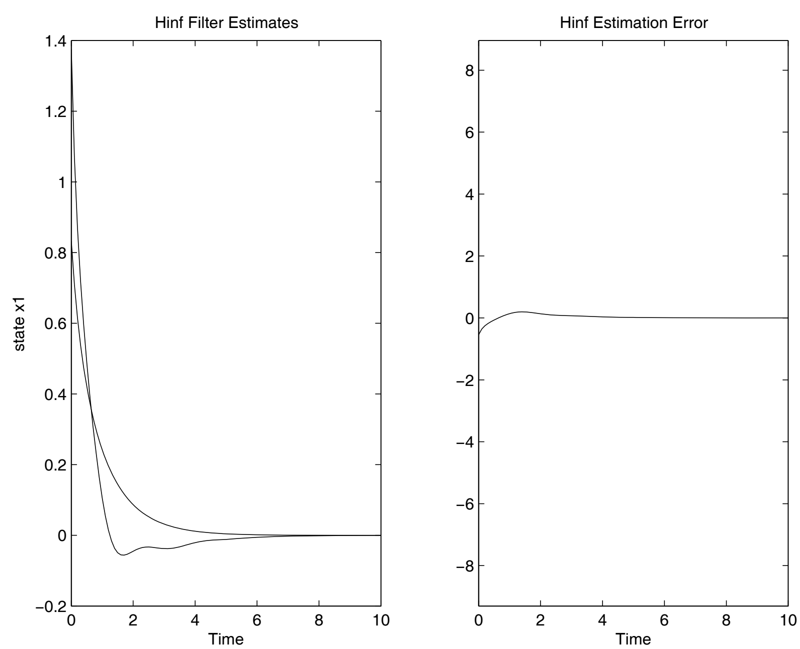

Figure 6.3 $\mathcal{H}_\infty$ *singular filter performance with unknown initial condition and* $\mathcal{L}_2$*-bounded disturbance*

get $\bar{x}_2 = -x_1$*, and substituting in (6.40), we get the reduced system*

$$\dot{x}_1 = -x_1^3 - x_1$$

which is locally asymptotically stable about $x = 0$*. Then, it can be checked that, the function* $\breve{V}(x) = \frac{1}{2}\breve{x}_1^2$ *solves the HJIE (6.34), and consequently, we have the filter gain*

$$\breve{L}(\breve{x}_1, y) = -\frac{y - \breve{x}_1}{\breve{x}_1},$$

where again $\breve{L}(\breve{x}_1, y)$ *is set equal to zero if* $|\breve{x}_1| < \epsilon$ *small. The result of the simulation is the same as for the singular filter above.*

6.3 $\mathcal{H}_\infty$ Filtering for Discrete-time Systems

In this section, we present the counterpart $\mathcal{H}_\infty$ filtering results for discrete-time nonlinear descriptor systems. We similarly present solutions using singular and normal filters, and examples are also presented to demonstrate the results. We begin with the problem definition and other preliminaries.

6.3.1 Problem Definition and Preliminaries

The general set-up for studying $\mathcal{H}_\infty$ filtering problems is shown in Fig. 6.4, where $\mathbf{P}_k$ is the plant, while $\mathbf{F}_k$ is the filter. The noise signal $w \in \mathcal{S}'$ is in general a bounded power signal (e.g. an ℓ_2 signal) which belongs to the set $\mathcal{P}'$ of bounded power signals, and $\tilde{z} \in \mathcal{P}'$, is also a bounded power signal. Thus, the induced norm from w to $\tilde{z}$ (the penalty variable to be defined later) is the ℓ_∞-norm of the interconnected system $\mathbf{F}_k \circ \mathbf{P}_k$ (where the operator "$\circ$" implies composition of input-output maps), i.e.,

$$\|\mathbf{F}_k \circ \mathbf{P}_k\|_{\ell_\infty} \triangleq \sup\nolimits_{0 \neq w \in \mathcal{P}'} \frac{\|\tilde{z}\|_{\mathcal{P}'}}{\|w\|_{\mathcal{P}'}}, \tag{6.43}$$

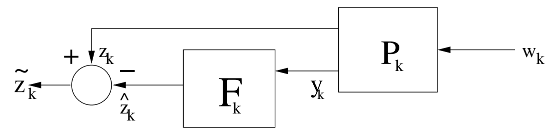

Figure 6.4 Set-up for Discrete-time $\mathcal{H}_\infty$ Filtering

$$\mathcal{P}' \triangleq \{w : \; w \in \ell_\infty, \; R_{ww}(k), \; S_{ww}(j\omega) \text{ exist for all } k \text{ and all } \omega \text{ resp.}, \|w\|_{\mathcal{P}'} < \infty\}$$

$$\|z\|^2_{\mathcal{P}'} \triangleq \lim_{K\to\infty} \frac{1}{2K} \sum_{k=-K}^{K} \|z_k\|^2,$$

and R_{ww}, $S_{ww}(j\omega)$ are the autocorrelation and power spectral density matrices of w. Notice also that, $\|(.)\|_{\mathcal{P}'}$ is a seminorm. In addition, if the plant-filter pair is stable, we replace the induced ℓ_∞-norm above by the equivalent $\mathcal{H}_\infty$ subspace norm.

At the outset, we consider the following affine nonlinear causal descriptor model of the plant which is defined on $\mathcal{X} \subseteq \Re^n$ with zero control input:

$$\mathbf{P}_D^{ad} \; : \; \begin{cases} Ex_{k+1} &= f(x_k) + g_1(x_k)w_k; \;\; x(k_0) = x^0 \\ y_k &= h_2(x_k) + k_{21}(x_k)w_k, \end{cases} \tag{6.44}$$

where $x \in \mathcal{X}$ is the semistate vector; $w \in \mathcal{W} \subset \Re^r$ is an unknown disturbance (or noise) signal, which belongs to the set $\mathcal{W}$ of admissible exogenous inputs; $y \in \mathcal{Y} \subset \Re^m$ is the measured output (or observation) of the system, and belongs to $\mathcal{Y}$, the set of admissible measured-outputs.

The functions $f : \mathcal{X} \to \mathcal{X}$, $g_1 : \mathcal{X} \to \mathcal{M}^{n_1 \times r}(\mathcal{X})$, where $\mathcal{M}^{i\times j}$ is the ring of $i \times j$ smooth matrices over $\mathcal{X}$, $h_2 : \mathcal{X} \to \Re^m$, and $k_{21} : \mathcal{X} \to \mathcal{M}^{m\times r}(\mathcal{X})$ are real C^∞ functions of x, while $E \in \Re^{n\times n}$ is a constant but singular matrix. Furthermore, we assume without any loss of generality that the system (6.44) has an isolated equilibrium-point at $x = 0$ which is admissible and is such that $f(0) = 0$, $h_2(0) = 0$. We also assume that there exists at least one solution $x(k, k_0, x^0, w)$ $\forall k \in \mathbf{Z}$ for the system, for all admissible initial conditions x^0, for all $w \in \mathcal{W}$. The initial condition x^0 is said to be admissible if the solution x_k is unique and impulse-free for all $k \in [k_0, \infty)$. For simplicity we also make the following assumptions on the plant.

Assumption 6.3.1. *The system matrices are such that*

$$\begin{aligned} k_{21}(x)g_1^T(x) &= 0, \\ k_{21}(x)k_{21}^T(x) &= I. \end{aligned}$$

In addition, the following standing assumptions will be made on the system.

Assumption 6.3.2. *Let* $A = \frac{\partial f}{\partial x}(\bar{x})$, $\bar{x} \in \mathcal{O} \subset \mathcal{X}$. *Then, the system (6.44) is locally admissible, implies the following hold:*

1. *the system is locally regular at each* $\bar{x} \in \mathcal{O}$ *and hence locally solvable, i.e,* $det(zE - A) \not\equiv 0$ *for all* $z \in \mathbf{C}$*;*

2. *the system is locally impulse-free at each* $\bar{x} \in \mathcal{O}$*, i.e.,* $deg(det(zE - A)) = rank(E)\ \forall z \in \mathbf{C}$*;*

3. *the system is locally asymptotically stable, i.e.,* (E, A) *is Hurwitz at* $\bar{x} = 0$*.*

The suboptimal $\mathcal{H}_\infty$ local filtering/state-estimation problem is defined as follows.

Definition 6.3.1. *(Suboptimal* $\mathcal{H}_\infty$ *Local State Estimation or Filtering Problem). Find a filter,* $\mathbf{F}_k$*, for estimating the state* x_k *or a function of it,* $z_k = h_1(x_k)$*, from observations* $\mathsf{Y}_k \triangleq \{y(i) : i \leq k\}$*, of* $y(i)$ *up to time* k*, to obtain the estimate*

$$\hat{x}_k = \mathbf{F}_k(\mathsf{Y}_k),$$

such that, the ℓ_2*-gain from the input* w *to some suitable penalty function* $\tilde{z}$ *is rendered less or equal to some desired number* $\gamma > 0$*, i.e.*

$$\sum_{k=k_0}^{\infty} \|\tilde{z}_k\|^2 \leq \gamma^2 \sum_{k=k_0}^{\infty} \|w_k\|^2 \tag{6.45}$$

for all admissible initial conditions $x^0 \in \mathcal{O} \subset \mathcal{X}$ *for all* $w \in \ell_2[k_0, \infty)$*. Moreover, if the filter solves the problem for all admissible* $x^0 \in \mathcal{X}$*, we say the problem is solved globally.*

We shall adopt the following definition of local zero-input observability which we coined from (Ozcaldiran, 1992), (Vidyasagar, 1993).

Definition 6.3.2. *For the nonlinear system* $\mathbf{P}_D^{da}$*, we say that, it is locally weakly zero-input observable, if for all states* $x_1,\ x_2 \in U \subset \mathcal{X}$ *and input* $w(.) = 0,\ k > k_0$

$$y(k; Ex_1(k_0-), w) \equiv y(k; Ex_2(k_0-), w) \Longrightarrow Ex_1(k_0) = Ex_2(k_0); \tag{6.46}$$

the system is said to be locally zero-input observable if

$$y(k; Ex_1(k_0-), w) \equiv y(k; Ex_2(k_0-), w) \Longrightarrow x_1(k_0) = x_2(k_0); \tag{6.47}$$

where $y(., Ex_i(k_0-), w),\ \ i = 1, 2$ *is the output of the system with the initial condition* $Ex_i(t_0-)$*; and the system is said to be locally strongly zero-input observable if*

$$y(k; Ex_1(k_0-), w) \equiv y(k; Ex_2(k_0-), w) \Longrightarrow x_1(k_0-) = x_2(t_0-). \tag{6.48}$$

Moreover, the system is said to be globally (weakly, strongly) zero-input observable, if it is locally (weakly, strongly)-observable at each $x(k_0) \in \mathcal{X}$ *or* $U = \mathcal{X}$*.*

In the sequel, we shall not distinguish between local observability and strong local observability. Moreover, in the next two subsections, we discuss singular and normal filters for the $\mathcal{H}_\infty$ state estimation problem defined above. For this purpose, we assume throughout that the noise signal $w \in \mathcal{W} \subset \ell_2[k_0, \infty)$.

6.3.2 Discrete-time $\mathcal{H}_\infty$ Singular Filters

In this subsection, we discuss full-order $\mathcal{H}_\infty$ singular filters for the system in the usual Kalman-Luenberger type structure:

$$\mathbf{F}^{ad}_{DS1}\begin{cases} E\hat{x}_{k+1} &= f(\hat{x}_k)+g_1(\hat{x}_k)\hat{w}^\star_k+L(\hat{x}_k,y_k)[y_k-h_2(\hat{x}_k-k_{21}(\hat{x}_k)\hat{w}^\star_k],\ \ \hat{x}(k_0)=0 \\ \tilde{z}_k &= y_k-h_2(\hat{x}_k) \end{cases} \tag{6.49}$$

where $\hat{x}\in\mathcal{X}$ is the filter state and $\hat{L}:\mathcal{X}\times\mathcal{Y}\to\Re^{n\times m}$ is the gain matrix of the filter, $\tilde{z}\in\Re^m$ is the penalty variable or estimation error, and $\hat{w}^\star$ is the estimated certainty-equivalent worst system noise. Notice also here that, we are not including the term $k_{12}w$ in the filter design, because its effect on $\hat{w}^\star$ is negligible, and to simplify the presentation.

The problem can then be formulated as a dynamic optimization problem with the following cost functional:

$$\begin{aligned}\min_{\hat{L}\in\Re^{n\times m}}\sup\nolimits_{w\in\mathcal{W}}J(\hat{L},w) &= \frac{1}{2}\sum_{k=k_0}^{\infty}[\|\tilde{z}_k\|^2-\gamma^2\|w_k\|^2], \quad s.t.\ (6.49), \\ &\quad \text{and with } w=0, \quad \lim_{k\to\infty}\{\hat{x}_k-x_k\}=0. \end{aligned}\tag{6.50}$$

To solve the above problem, we form the Hamiltonian function[2] $H:\mathcal{X}\times\mathcal{Y}\times\Re^{n\times m}\times\Re\to\Re$ (Wang, 2008):

$$\begin{aligned}H(\hat{x},y,w,L,V) &= V(E(f(\hat{x})+g_1(\hat{x})w+L(\hat{x},y)(y-h_2(\hat{x}))-k_{21}(\hat{x})\hat{w}^\star),y)- \\ &\quad V(E\hat{x},y_{k-1})+\frac{1}{2}(\|\tilde{z}\|^2-\gamma^2\|w\|^2)\end{aligned}\tag{6.51}$$

for some C^2 function $V:\mathcal{X}\times\mathcal{Y}\to\Re$ and where $x=x_k$, $y=y_k$, $\tilde{z}=\{\tilde{z}_k\}$, $w=\{w_k\}$. Notice also here that, we are only using y_{k-1} in the above expression (6.51) to distinguish between $y_k=y$ and y_{k-1}. Otherwise, (6.51) holds for all y and is C^1 in all its arguments. Then, the optimal gain $L^\star$ can be obtained by minimizing H with respect to L in the above expression

[2] Our definition is slightly different from Reference (Wang, 2008) in order to maintain the symmetry of the Hamiltonian

(6.51), as

$$\begin{aligned}
\hat{w}^\star &= \frac{1}{\gamma^2} g_1^T(\hat{x}) \frac{\partial V(\lambda, y)}{\partial \lambda}\bigg|_{\lambda = E(f(\hat{x}) + g_1(\hat{x})w + L(\hat{x},y)(y - h_2(\hat{x}) - k_{21}(\hat{x})w)), w = w^\star} && (6.52)\\
L^\star &= \arg\min_L H(\hat{x}, w^\star, y, L, V). && (6.53)
\end{aligned}$$

where $\partial V/\partial\lambda$ is the row vector of first-order partial derivatives of V with respect to λ. Because the Hamiltonian function (6.51) is not a linear or quadratic function of the gain L, only implicit solutions can be obtained by solving the above equations. Thus, the only way to obtain an explicit solution, is to use an approximate scheme. Accordingly, consider a first-order Taylor approximation of the Hamiltonian function (6.51) about $(Ef(\hat{x}), y)$ and in the direction of the estimator state vector $E\hat{x}$, denoted by $\widehat{H}$:

$$\begin{aligned}
\widehat{H}(\hat{x}, y, w, \hat{L}, \hat{V}) &= \hat{V}(Ef(\hat{x}), y) + \hat{V}_{E\hat{x}}(Ef(\hat{x}), y)[E(f(\hat{x}) + g_1(\hat{x})w + \hat{L}(\hat{x}, y)(y - h_2(\hat{x}) - \\
&\quad k_{21}(\hat{x})w)] - \hat{V}(E\hat{x}, y_{k-1}) + \frac{1}{2}(\|\tilde{z}\|^2 - \gamma^2\|w\|^2) + O(\|\hat{x}\|^2), && (6.54)
\end{aligned}$$

where $\hat{V}$, $\hat{L}$ are the corresponding approximate functions. Applying the necessary condition for the worst-case noise, we get

$$\frac{\partial \widehat{H}}{\partial w}\bigg|_{w=\hat{w}^\star} = 0 \Longrightarrow \hat{w}^\star = \frac{1}{\gamma^2}[g_1^T(\hat{x}) - k_{21}^T(\hat{x})\hat{L}^T(\hat{x}, y)]E\hat{V}_{E\hat{x}}^T(Ef(\hat{x}), y). \quad (6.55)$$

Substituting $\hat{w}^\star$ in (6.54) and completing the squares in $E\hat{L}$, we get

$$\begin{aligned}
\widehat{H}(\hat{x}, y, \hat{w}^\star, \hat{L}, \hat{V}) &= \hat{V}(Ef(\hat{x}), y) + \hat{V}_{E\hat{x}}(Ef(\hat{x}), y)Ef(\hat{x}) - \hat{V}(E\hat{x}, y_{k-1}) + \\
&\quad \frac{1}{2\gamma^2}\left\|\hat{L}^T(\hat{x}, y)E^T\hat{V}_{E\hat{x}}(Ef(\hat{x}), y) + \gamma^2(y - h_2(\hat{x}))\right\|^2 + \\
&\quad \frac{1}{2\gamma^2}\hat{V}_{E\hat{x}}(Ef(\hat{x}), y)Eg_1(\hat{x})(\hat{x})g_1^T(\hat{x})E^T\hat{V}_{E\hat{x}}^T(Ef(\hat{x}), y) + \\
&\quad \frac{(1-\gamma^2)}{2}\|\tilde{z}\|^2. && (6.56)
\end{aligned}$$

Thus, setting

$$\hat{V}_{E\hat{x}}(Ef(\hat{x}), y)E\hat{L}^\star(\hat{x}, y) = -\gamma^2(y - h_2(\hat{x}))^T \quad (6.57)$$

minimizes $\widehat{H}(.,.,\hat{w}^\star,\hat{L},.)$ and guarantees that the saddle-point condition (Basar, 1982)

$$\widehat{H}(\hat{x},y,\hat{w}^\star,\hat{L}^\star,\hat{V}) \leq \widehat{H}(\hat{x},y,\hat{w}^\star,\hat{L},\hat{V}) \quad \forall \hat{L} \in \Re^{n\times m} \tag{6.58}$$

is satisfied. Finally, substituting the above expression for $\hat{L}^\star$ in (6.56) and setting

$$\widehat{H}(\hat{x},y,\hat{w}^\star,\hat{L}^\star,\hat{V}) = 0$$

yields the following DHJIE:

$$\begin{aligned}\hat{V}(Ef(\hat{x}),y) + \hat{V}_{E\hat{x}}(Ef(\hat{x}),y)Ef(\hat{x}) + \tfrac{1}{2\gamma^2}\hat{V}_{E\hat{x}}(Ef(\hat{x}),y)Eg_1(\hat{x})(\hat{x})g_1^T(\hat{x})E^T\hat{V}_{E\hat{x}}^T(Ef(\hat{x}),y) - \\ \hat{V}(E\hat{x},y_{k-1}) + \tfrac{(1-\gamma^2)}{2}(y-h_2(\hat{x}))^T(y-h_2(\hat{x})) = 0, \quad \hat{V}(0,0) = 0.\end{aligned} \tag{6.59}$$

Moreover, from (6.54), (6.59), we can write

$$\widehat{H}(\hat{x},y,\hat{w},\hat{L}^\star,\hat{V}) = \widehat{H}(\hat{x},y,\hat{w}^\star,\hat{L}^\star,\hat{V}) - \frac{1}{2}\gamma^2\|w-\hat{w}^\star\|^2, \tag{6.60}$$

and hence,

$$\widehat{H}(\hat{x},y,\hat{w},\hat{L}^\star,\hat{V}) \leq \widehat{H}(\hat{x},y,\hat{w}^\star,\hat{L}^\star,\hat{V}). \tag{6.61}$$

Thus, combining (6.58) and (6.61), we see that the saddle-point conditions (Basar, 1982)

$$\widehat{H}(\hat{x},y,\hat{w},\hat{L}^\star,\hat{V}) \leq \widehat{H}(\hat{x},y,\hat{w}^\star,\hat{L}^\star,\hat{V}) \leq \widehat{H}(\hat{x},y,\hat{w}^\star,\hat{L},\hat{V}) \tag{6.62}$$

are satisfied, and the pair $(\hat{w}^\star,\hat{L}^\star)$ constitutes a saddle-point solution to the dynamic game (6.50), (6.44). Consequently, we have the following result.

Proposition 6.3.1. *Consider the nonlinear system (6.44) and the $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\mathbf{P}_\mathbf{D}^\mathbf{ad}$ satisfies Assumptions 6.3.1, locally asymptotically stable about the equilibrium-point $x = 0$, and zero-input observable. Further, suppose for some $\gamma > 0$, there exist a C^1 positive-semidefinite function $\hat{V} : \hat{N} \times \hat{\Upsilon} \to \Re_+$ locally defined in a neighborhood $\hat{N} \times \hat{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ of the origin $(\hat{x},y) = (0,0)$, and a matrix function*

$\hat{L} : \hat{N} \times \hat{\Upsilon} \to \Re^{n\times m}$, satisfying the DHJIE (6.59) together with the side-condition (6.57). Then, the filter $\mathbf{F}^{ad}_{DS1}$ solves the $\mathcal{H}_\infty$ filtering problem for the system locally in $\hat{N}$.

Proof: The optimality of the filter gain $\hat{L}^\star$ has already been shown above. It remains to prove that the ℓ_2-gain condition (6.45) is satisfied, and there is asymptotic convergence of the estimation error to zero. Accordingly, let $\hat{V}(E\hat{x}, y) \geq 0$ be a C^1 solution of the DHJIE (6.59). Then, taking the time variation of $\hat{V}$ along a trajectory of (6.49) with $\hat{L} = \hat{L}^\star$, we get

$$\begin{aligned}
\hat{V}(E\hat{x}_{k+1}, y) \;&\approx\; \hat{V}(Ef(\hat{x}), y) + \hat{V}_{E\hat{x}}(Ef(\hat{x}), y)[E(f(\hat{x}) + g_1(\hat{x})w + \\
&\quad\; \hat{L}^\star(\hat{x}, y)(y - h_2(\hat{x}) - k_{21}(\hat{x})w)] \\
&= \Big\{\hat{V}(Ef(\hat{x}), y) + \hat{V}_{E\hat{x}}(Ef(\hat{x}), y)Ef(\hat{x}) + \frac{(1-\gamma^2)}{2}\|\tilde{z}\|^2 + \\
&\quad\; \frac{1}{2\gamma^2}\hat{V}_{E\hat{x}}(Ef(\hat{x}), y)Eg_1(\hat{x})(\hat{x})g_1^T(\hat{x})E^T\hat{V}^T_{E\hat{x}}(Ef(\hat{x}), y)\Big\} - \\
&\quad\; \frac{1}{2}\gamma^2\|w - \hat{w}^\star\|^2 + \frac{1}{2}\gamma^2\|w\|^2 - \frac{1}{2}\|\tilde{z}\|^2 \\
&\leq \hat{V}(E\hat{x}, y_{k-1}) + \frac{1}{2}\gamma^2\|w\|^2 - \frac{1}{2}\|\tilde{z}\|^2, \qquad (6.63)
\end{aligned}$$

where use has been made of the first-order Taylor approximation above, and the last inequality (6.63) follows from using the DHJIE (6.59). Summing the above inequality from $k = k_0$ to ∞ we get that

$$\hat{V}(x_\infty, y_\infty) - \hat{V}(k_0, y_{k_0-1}) \leq \frac{1}{2}\sum_{k=k_0}^{\infty}(\gamma^2\|w_k\|^2 - \|\tilde{z}_k\|^2),$$

and therefore the ℓ_2-gain condition (6.45) is satisfied. In addition, setting $w = 0$ in the inequality (6.63), we have

$$\hat{V}(E\hat{x}_{k+1}, y_k) - \hat{V}(E\hat{x}_k, y_{k-1}) \leq -\frac{1}{2}\|\tilde{z}_k\|^2,$$

and by Lyapunov's theorem, the filter dynamics is stable, i.e., $V(E\hat{x}, y)$ is non-increasing along a trajectory of (6.49). Further, the condition that $\hat{V}(E\hat{x}_{k+1}, y_k) \equiv \hat{V}(E\hat{x}, y_{k-1})$ $\forall k \geq k_s$, for some k_s, implies that $\tilde{z}_k \equiv 0$, which further implies that $y_k = h_2(\hat{x}_k)$ $\forall k \geq k_s$. By the zero-input observability of the system, this implies that $\hat{x}_k = x_k \; \forall k \geq k_s$. $\square$

The result of the theorem can be specialized to the linear descriptor system

$$\mathbf{P}_D^{dl} : \begin{cases} Ex_{k+1} &= Ax_k + B_1 w_k; \quad x(k_0) = x^0 \\ y_k &= C_2 x_k + D_{21} w_k, \quad D_{21}^T B_1 = 0, \quad D_{21} D_{21}^T = I, \end{cases} \tag{6.64}$$

where $E \in \Re^{n\times n}$, $A \in \Re^{n\times n}$, $B_1 \in \Re^{n\times r}$, $C_2 \in \Re^{m\times n}$, $D_{21} \in \Re^{m\times r}$.

Corollary 6.3.1. *Consider the linear descriptor system (6.64) and the $\mathcal{H}_\infty$ filtering problem for this system. Suppose the plant $\mathbf{P}_D^{dl}$ is locally asymptotically stable about the equilibrium-point $x = 0$ and observable. Suppose further, there exist symmetric positive-semidefinite matrices $\hat{P} \in \Re^{n\times n}$, $\hat{Q} \in \Re^{m\times m}$, and a matrix $\hat{L} \in \Re^{n\times m}$, satisfying the linear matrix-inequalities (LMIs)*

$$\begin{bmatrix} A^T E^T \hat{P} E A - E^T \hat{P} E + (1-\gamma^2) C_2^T C_2 & A^T E^T \hat{P} E B_1 & (1-\gamma^2) C_2^T & 0 \\ B_1^T E P E A & -\gamma^2 I & 0 & 0 \\ (1-\gamma^2) C_2 & 0 & \hat{Q} - I & 0 \\ 0 & 0 & 0 & -\hat{Q} \end{bmatrix} \leq 0 \tag{6.65}$$

$$\begin{bmatrix} 0 & \frac{1}{2}(A^T E^T \hat{P} E L - \gamma^2 C_2^T) \\ \frac{1}{2}(A^T E^T \hat{P} E L - \gamma^2 C_2^T)^T & (\gamma^2 - \delta_1) I \end{bmatrix} \leq 0 \tag{6.66}$$

for some number $\delta_1 \geq 1$. Then the filter

$$\mathbf{F}_{DS1}^{dl} : \begin{cases} E\hat{x}_{k+1} &= (A + \frac{1}{\gamma^2} B_1 B_1^T E P E A + \frac{1}{\gamma^2} \hat{L} \hat{L}^T E P E A)\hat{x}_k + \hat{L}(y - C_2 \hat{x}_k); \\ & \hat{x}(k_0) = 0 \\ \hat{z}_k &= y_k - C_2 \hat{x}_k \end{cases} \tag{6.67}$$

solves the $\mathcal{H}_\infty$ filtering problem for the system.

Proof: Take $\hat{V}(E\hat{x}, y) = \frac{1}{2}(\hat{x}^T E^T \hat{P} E \hat{x} + y^T \hat{Q} y)$, $\quad \hat{P} > 0$ and apply the result of the Proposition. $\square$

Notice similarly however, since the system is inherently constrained, convergence of the estimates may be slow with filter $\mathbf{F}_{DS1}^{ad}$. Therefore, to guarantee better convergence, we propose a proportional-integral (PI) filter configuration (Gao, 2004), (Koenig, 1995) to further improve the convergence. Thus, we consider the following class of filters:

$$\mathbf{F}^{ad}_{DS2}\begin{cases} E\check{x}_{k+1} &= f(\check{x}_k)+g_1(\check{x})\check{w}^\star_k+\check{L}_1(\check{x}_k,\xi_k,y_k)(y_k-h_2(\check{x}_k)-k_{21}(\check{x}_k)\check{w}^\star_k)+ \\ & \quad \check{L}_2(\check{x}_k,\xi_k,y_k)\xi_k, \quad \check{x}(k_0)=0 \\ \xi_{k+1} &= y_k-h_2(\check{x}_k) \\ \check{z}_k &= y_k-h_2(\check{x}_k) \end{cases} \tag{6.68}$$

where $\check{x}\in\mathcal{X}$ is the filter state, $\check{w}^\star$ is the estimated worst-case noise of the system, $\xi\in\Re^m$ is the integrator state, and $\check{L}_1,\check{L}_2:\mathcal{X}\times\mathcal{Y}\times\mathcal{Y}\to\Re^{n\times m}$ are the proportional and integral gain matrices of the filter respectively. Similarly, using manipulations as in Proposition 6.3.1, we can arrive at the following result.

Theorem 6.3.1. *Consider the nonlinear system (6.44) and the $\mathcal{H}_\infty$ local filtering problem for this system. Suppose the plant $\mathbf{P^{ad}_D}$ satisfies Assumption 6.3.1, is locally asymptotically stable about the equilibrium-point $x=0$, and locally zero-input observable. Further, suppose there exist a C^2 positive-definite function $\check{V}:\check{N}\times\check{\Xi}\times\check{\Upsilon}\to\Re_+$ locally defined in a neighborhood $\check{N}\times\check{\Xi}\times\hat{\Upsilon}\subset\mathcal{X}\times\mathcal{Y}\times\mathcal{Y}$ of the origin $(\check{x},\xi,y)=(0,0,0)$, and matrix functions $\check{L}_1,\check{L}_2:\check{N}\times\check{\Xi}\times\check{\Upsilon}\to\Re^{n\times m}$, satisfying the DHJIE*

$$\begin{aligned} &\check{V}(Ef(\check{x}),\xi,y)+\check{V}_{E\check{x}}(Ef(\check{x}),\xi,y)Ef(\check{x})+ \\ &\tfrac{1}{2\gamma^2}\hat{V}_{E\check{x}}(Ef(\check{x}),y)Eg_1(\check{x})(\check{x})g_1^T(\check{x})E^T\check{V}^T_{E\check{x}}(Ef(\check{x}),y)- \\ &\check{V}(E\check{x},\xi,y_{k-1})+\check{V}_\xi(Ef(\check{x}),\xi,y)(y-h_2(\check{x}))-\xi^T\xi+ \\ &\tfrac{(1-\gamma^2)}{2}(y-h_2(\check{x}))^T(y-h_2(\check{x}))=0, \qquad \check{V}(0,0,0)=0, \end{aligned} \tag{6.69}$$

together with the side-conditions

$$\check{V}_{E\check{x}}(E\check{x},\xi,y)E\check{L}^\star_1(\check{x},\xi,y) = -\gamma^2(y-h_2(\check{x}))^T, \tag{6.70}$$

$$\check{V}_{E\check{x}}(E\check{x},\xi,y)E\check{L}^\star_2(\check{x},\xi,y) = -\xi^T. \tag{6.71}$$

Then, the filter $\mathbf{F}^{ad}_{DS2}$ solves the $\mathcal{H}_\infty$ local filtering problem for the system locally in $\check{N}$.

Proof: The proof follows along similar lines as Proposition 6.3.1. □

In the next section, we consider the design of normal filters for the system.

6.3.3 Discrete-time $\mathcal{H}_\infty$ Normal Filters

In this subsection, we discuss normal filters for the system (6.44). We shall consider the design of both full-order and reduced-order filters. We start with the full-order filter first, and in this regard, without any loss of generality, we can assume that E is of the form

$$E = \begin{pmatrix} I_{q\times q} & 0 \\ 0 & 0 \end{pmatrix}.$$

This follows from matrix theory and can easily be proven using the singular-value decomposition (SVD) of E. It follows that, the system can be represented in the canonical form of a differential-algebraic system

$$\bar{\mathbf{P}}_D^{ad} : \begin{cases} x_{1,k+1} &= f_1(x_k) + g_{11}(x_k)w_k; \;\; x(k_0) = x^0 \\ 0 &= f_2(x_k) + g_{21}(x_k)w_k \\ y &= h_2(x_k) + k_{21}(x_k)w_k \end{cases} \tag{6.72}$$

where $dim(x_1) = q$, $f_1(0) = 0$, $f_2(0) = 0$. We also assume the following counterpart of Assumption 6.3.1 for simplicity.

Assumption 6.3.3. *The system matrices in (6.72) are such that*

$$\begin{aligned} k_{21}(x)[g_{11}^T(x) \;\; g_{21}(x)] &= 0, \\ k_{21}(x)k_{21}^T(x) &= I. \end{aligned}$$

Then, if we define

$$x_{2,k+1} = f_2(x_k) + g_{21}(x_k)w_k,$$

where $x_{2,k+1}$ is a fictitious state vector, and $dim(x_2) = n - q$, the system (6.72) can be represented by a normal state-space system as

$$\widetilde{\mathbf{P}}^{ad}_D : \begin{cases} x_{1,k+1} &= f_1(x_k) + g_{11}(x_k)w_k; \quad x_1(k_0) = x^{10} \\ x_{2,k+1} &= f_2(x_k) + g_{21}(x_k)w_k; \quad x_2(k_0) = x^{20} \\ y &= h_2(x_k) + k_{21}(x_k)w_k. \end{cases} \tag{6.73}$$

Now define the following set $\Omega_o \subset \mathcal{X}$

$$\Omega_o = \{(x_1, x_2) \in \mathcal{X} \mid x_{2,k+1} \equiv 0, \quad k = 1, \ldots\}. \tag{6.74}$$

Then, we have the following system equivalence

$$\widetilde{\mathbf{P}}^{ad}_D|_{\Omega_o} = \bar{\mathbf{P}}^{ad}_D. \tag{6.75}$$

Thus, to estimate the states of the system (6.72), we need to stabilize the system (6.73) about Ω_o and then design a filter for the resulting system. For this purpose, we consider the following class of filters

$$\mathbf{F}^{ad}_{DN3} \begin{cases} \grave{x}_{1,k+1} &= f_1(\grave{x}_k) + g_{11}(\grave{x}_k)\grave{w}^\star_k + \grave{L}_1(\grave{x}_k, y_k)[y_k - h_2(\grave{x}_k) - k_{21}(\grave{x}_k)\grave{w}^\star_k], \\ & \quad \grave{x}_1(k_0) = 0 \\ \grave{x}_{2,k+1} &= f_2(\grave{x}_k) + g_{21}(\grave{x}_k)\grave{w}^\star_k + g_{22}(x_k)\alpha_2(\grave{x}_k) + \grave{L}_2(\grave{x}_k, y_k)[y_k - \\ & \quad h_2(\grave{x}_k) - k_{21}(\grave{x}_k)\grave{w}^\star_k], \quad \grave{x}_2(k_0) = 0 \\ \grave{z}_k &= y_k - h_2(\grave{x}_k), \end{cases} \tag{6.76}$$

where $\grave{x} \in \mathcal{X}$ is the filter state, $\grave{w}^\star$ is the worst-case estimated system noise, $\grave{L}_1 : \mathcal{X} \times \mathcal{Y} \to \Re^{q\times m}$, $\grave{L}_2 : \mathcal{X} \times \mathcal{Y} \to \Re^{(n-q)\times m}$ are the filter gain matrices, and $\tilde{g}_{22} : \mathcal{X} \to \mathcal{M}^{(n-q)\times p}$ is a gain matrix for the artificial control input $u = \alpha_2(\grave{x}) \in \Re^p$ required to stabilize the corresponding filter dynamics $\grave{x}_{2,k+1}$ about

$$\grave{\Omega}_o = \{(\grave{x}_1, \grave{x}_2) \in \mathcal{X} \mid \grave{x}_{2,k+1} \equiv 0, \quad k = 1, \ldots\}. \tag{6.77}$$

Accordingly, we make the following assumption.

Assumption 6.3.4. *The pair* $\{f_2, \tilde{g}_{22}\}$ *is locally stabilizable, i.e.,* $\exists$ *a control law* $\alpha_2(\grave{x}_2)$ *and a Lyapunov-function (LF),* $\bar{V} > 0$*, such that* $\bar{V}(f_2(\grave{x}) - \tilde{g}_{22}(\grave{x})\alpha_2(\grave{x})) - \bar{V}(\grave{x}) < 0 \; \forall \grave{x} \in \grave{N} \subset \mathcal{X}$.

Thus, if Assumption 6.3.4 holds, then we can make $\alpha_2 = \alpha_2(\grave{x}, \varepsilon)$, where $\varepsilon > 0$ is small, a high-gain feedback (Young, 1977) to constrain the dynamics on Ω_o as fast as possible. Then, we proceed to design the gain matrices $\grave{L}_1$, $\grave{L}_2$ to estimate the states using similar approximations as in the previous section. Using the first-order Taylor approximation, we have the following result.

Proposition 6.3.2. *Consider the nonlinear system (6.72) and the* $\mathcal{H}_\infty$ *local filtering problem for this system. Suppose the plant* $\bar{\mathbf{P}}^{\mathbf{ad}}_{\mathbf{D}}$ *satisfies Assumptions 6.3.3, 6.3.4, is locally asymptotically stable about the equilibrium-point* $x = 0$*, and zero-input observable. Further, suppose there exist a* C^1 *positive-semidefinite function* $\grave{V} : \grave{N} \times \grave{\Upsilon} \to \Re_+$*, locally defined in a neighborhood* $\grave{N} \times \grave{\Upsilon} \subset \mathcal{X} \times \mathcal{Y}$ *of the origin* $(\grave{x}_1, \grave{x}_2, y) = (0, 0, 0)$*, and matrix functions* $\grave{L}_1 : \grave{N} \times \grave{\Upsilon} \to \Re^{q \times m}$, $\grave{L}_2 : \grave{N} \times \grave{\Upsilon} \to \Re^{n-q \times m}$*, satisfying the DHJIE:*

$$
\begin{gathered}
\grave{V}(f_1(\grave{x}), f_2(\grave{x}), y) - \grave{V}(\grave{x}_1, \grave{x}_2, y_{k-1}) + \\
\tfrac{1}{2\gamma^2} \hat{V}_{\grave{x}_1}(f_1(\grave{x}), f_2(\grave{x}), y) g_{11}(\grave{x})(\grave{x}) g_{11}^T(\grave{x}) \grave{V}_{\grave{x}_1}^T(f_1(\grave{x}), f_2(\grave{x}), y) + \\
\tfrac{1}{\gamma^2} \grave{V}_{\grave{x}_1}(f_1(\grave{x}), f_2(\grave{x}), y) g_{11}(\grave{x})(\grave{x}) g_{12}^T(\grave{x}) \grave{V}_{\grave{x}_2}^T(f_1(\grave{x}), f_2(\grave{x}), y) - \\
\grave{V}_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y) g_{22}(\grave{x}) \alpha_2(\grave{x}, \varepsilon) + \\
\tfrac{1}{2\gamma^2} \grave{V}_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y) g_{21}(\grave{x})(\grave{x}) g_{21}^T(\grave{x}) \grave{V}_{\grave{x}_2}^T(f_1(\grave{x}), f_2(\grave{x}), y) + \\
\tfrac{(1-\gamma^2)}{2}(y - h_2(\grave{x}))^T (y - h_2(\grave{x})) = 0, \quad \grave{V}(0,0,0) = 0,
\end{gathered}
\tag{6.78}
$$

together with the side-conditions

$$
\grave{V}_{\grave{x}_1}(f_1(\grave{x}), f_2(\grave{x}), y) \grave{L}_1(\grave{x}_1, \grave{x}_2, y) + \grave{V}_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y) \grave{L}_2(\grave{x}_1, \grave{x}_2, y) = -\gamma^2 (y - h_2(\grave{x}))^T. \tag{6.79}
$$

Then, the filter $\mathbf{F}^{da}_{DN3}$ *solves the* $\mathcal{H}_\infty$*-filtering problem for the system locally in* $\grave{N}$.

Proof: Follows along same lines as Proposition 6.3.1.

A common DHJIE-Lyapunov function for both the stabilization (Guillard, 1996) and filter design can also be utilized in the above design procedure. This can be achieved optimally if we take

$$\begin{aligned} \alpha_2(\grave{x},\varepsilon) &= -\tfrac{1}{\varepsilon} g_{22}^T(\grave{x}) \bar{V}_{\grave{x}_2}^T(f_1(\grave{x}), f_2(\grave{x}), y) \\ \bar{V}_{\grave{x}_1}(f_1(\grave{x}), f_2(\grave{x}), y)\grave{L}_1(\grave{x}, y) + \bar{V}_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y)\grave{L}_2(\grave{x}, y) &= -\gamma^2 (y - h_2(\grave{x}))^T, \end{aligned}$$

where $\bar{V} \geq 0$ is a C^1 solution of the following DHJIE

$$\begin{gathered} \bar{V}(f_1(\grave{x}), f_2(\grave{x}), y) - \bar{V}(\grave{x}_1, \grave{x}_2, y_{k-1}) + \\ \tfrac{1}{2\gamma^2} \bar{V}_{\grave{x}_1}(f_1(\grave{x}), f_2(\grave{x}), y) g_{11}(\grave{x})(\grave{x}) g_{11}^T(\grave{x}) \bar{V}_{\grave{x}_1}^T(f_1(\grave{x}), f_2(\grave{x}), y) + \\ \tfrac{1}{\gamma^2} \bar{V}_{\grave{x}_1}(f_1(\grave{x}), f_2(\grave{x}), y) g_{11}(\grave{x})(\grave{x}) g_{12}^T(\grave{x}) \bar{V}_{\grave{x}_2}^T(f_1(\grave{x}), f_2(\grave{x}), y) - \\ \tfrac{1}{\varepsilon} \bar{V}_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y) g_{22}(\grave{x}) g_{22}^T(\grave{x}) \bar{V}_{\grave{x}_2}^T(f_1(\grave{x}), f_2(\grave{x}), y) + \\ \tfrac{1}{2\gamma^2} \bar{V}_{\grave{x}_2}(f_1(\grave{x}), f_2(\grave{x}), y) g_{21}(\grave{x})(\grave{x}) g_{21}^T(\grave{x}) \bar{V}_{\grave{x}_2}^T(f_1(\grave{x}), f_2(\grave{x}), y) + \\ \tfrac{(1-\gamma^2)}{2} (y - h_2(\grave{x}))^T (y - h_2(\grave{x})) = 0, \quad \bar{V}(0,0,0) = 0. \end{gathered} \tag{6.80}$$

Similarly, a normal PI-filter for the system (6.72) can also be designed. However, next we consider a reduced-order normal filter design. Accordingly, partition the state-vector x comformably with $rank(E) = q$ as $x = (x_1^T \;\; x_2^T)^T$ with $dim(x_1) = q$, $dim(x_2) = n - q$ and the state equations as

$$\breve{\mathbf{P}}_D^{ad} : \begin{cases} x_{1,k+1} &= f_1(x_{1,k}, x_{2,k}) + g_{11}(x_{1,k}, x_{2,k}) w_k; \quad x_1(k_0) = x^{10} \\ 0 &= f_2(x_{1,k}, x_{2,k}) + g_{21}(x_{1,k}, x_{2,k}) w_k; \quad x_2(k_0) = x^{20} \\ y_k &= h_2(x_k) + k_{21}(x_k) w_k \end{cases} \tag{6.81}$$

Then we make the following assumption.

Assumption 6.3.5. *The system is in the standard-form, i.e., the Jacobian matrix* $f_{2,x_2}(x_1, x_2)$ *is nonsingular in an open neighborhood* $\widetilde{U}$ *of* $(0,0)$ *and* $g_{21}(0,0) \neq 0$.

If Assumption 6.3.5 holds, then by the Implicit-function Theorem (Sastry, 1999), there exists a unique C^1 function $\varphi : \Re^q \times \mathcal{W} \to \Re^{n-q}$ and a solution

$$\bar{x}_2 = \phi(x_1, w)$$

to equation (6.81b). Thus, the system can be locally represented in $\widetilde{U}$ as the reduced-order system

$$\bar{\mathbf{P}}^{ad}_{rD} : \begin{cases} x_{1,k+1} &= f_1(x_{1,k}, \phi(x_{1,k}, w_k)) + g_{11}(x_{1,k}, \phi(x_{1,k}, w_k))w_k; \quad x_1(k_0) = x^{10} \\ y_k &= h_2(x_{1,k}, \phi(x_{1,k}, w_k)) + k_{21}(x_{1,k}, \phi(x_{1,k}, w_k))w_k. \end{cases} \tag{6.82}$$

We can then design a normal filter of the form

$$\mathbf{F}^{ad}_{DrN4} \begin{cases} \breve{x}_{1,k+1} &= f_1(\breve{x}_{1,k}, \phi(\breve{x}_{1,k}, \breve{w}^\star_{1,k})) + g_{11}(\breve{x}_{1,k}, \phi(\breve{x}_{1,k}, \breve{w}^\star_{1,k}))\breve{w}^\star_1 + \\ & \quad \breve{L}(\breve{x}_{1,k}, \phi(\breve{x}_{1,k}, \breve{w}^\star_{1,k}), y_k)[y - h_2(\breve{x}_{1,k}, \phi(\breve{x}_{1,k}, \breve{w}^\star)) - \\ & \quad k_{21}(\breve{x}_{1,k}, \phi(\breve{x}_{1,k}, \breve{w}^\star_1))\breve{w}^\star]; \\ & \quad \breve{x}_1(k_0) = 0 \\ \breve{z}_k &= y_k - h_2(\breve{x}_{1,k}, \phi(\breve{x}_{1,k}, \breve{w}^\star_1)), \end{cases} \tag{6.83}$$

where again all the variables have their corresponding previous meanings and dimensions. Consequently, we have the following result. Suppose that for simplicity, the following equivalent assumption is also satisfied by the subsystem (6.82).

Assumption 6.3.6. *The system matrices are such that*

$$\begin{aligned} k_{21}(x)g^T_{11}(x) &= 0, \\ k_{21}(x)k^T_{21}(x) &= I. \end{aligned}$$

Theorem 6.3.2. *Consider the nonlinear system (6.72) and the $\mathcal{H}_2$ local filtering problem for this system. Suppose for the plant $\bar{\mathbf{P}}^{\mathbf{ad}}_{\mathbf{D}}$ is locally asymptotically stable about the equilibrium-point $x = 0$, zero-input observable and Assumptions 6.3.5, 6.3.6 hold for the system. Further, suppose there exist a C^1 positive-semidefinite function $\breve{V} : \breve{N} \times \breve{\Upsilon} \to \Re_+$, locally defined in a neighborhood $\breve{N} \times \breve{\Upsilon} \subset \widetilde{U} \times \mathcal{Y}$ of the origin $(\breve{x}_1, y) = (0, 0)$, and a matrix function*

$\breve{L}:\breve{N}\times\breve{\Upsilon}\to\Re^{q\times m}$, *satisfying the DHJIE:*

$$\breve{V}(f_1(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star)),y)+\breve{V}_{\breve{x}_1}(f_1(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star)),y)f_1(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star))-\breve{V}(\breve{x}_1,y_{k-1})+$$
$$\tfrac{1}{2\gamma^2}\breve{V}_{\breve{x}_1}(f_1(\grave{x},\phi(\breve{x}_1,\breve{w}_1^\star)),y)g_{11}(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star))g_{11}^T(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star))\breve{V}_{\grave{x}_1}^T(f_1(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star)),y)$$
$$+\tfrac{(1-\gamma^2)}{2}(y-h_2(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star))^T(y-h_2(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star))=0,\quad \breve{V}(0,0)=0, \tag{6.84}$$

together with the side-condition

$$\breve{V}_{\breve{x}_1}(f_1(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star)),y)\breve{L}(\breve{x}_1,y)=-(y-h_2(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star)))^T, \tag{6.85}$$

$$\hat{w}_1^\star=\frac{1}{\gamma^2}[g_{11}^T(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star))-k_{21}^T(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star))\hat{L}^T(\hat{x}_1,y)]\breve{V}_{\grave{x}_1}^T(f_1(\breve{x}_1,\phi(\breve{x}_1,\breve{w}_1^\star)),y).$$

Then, the filter $\mathbf{F}_{DrN4}^{ad}$ *solves the* $\mathcal{H}_\infty$ *local filtering problem for the system in* $\breve{N}$.

Proof: Follows along same lines as Proposition 6.3.1.

Similarly, we can specialize the result of Theorem 6.3.2 to the linear system (6.64). The system can be rewritten in the form (6.72) as

$$\mathbf{P}_D^{ld}:\begin{cases} x_{k+1} &= A_1x_{1,k}+A_{12}x_{2,k}+B_{11}w_k;\quad x_1(k_0)=x^{10}\\ 0 &= A_{21}x_{1,k}+A_2x_{2,k}+B_{21}w_k;\quad x_2(k_0)=x^{20}\\ y_k &= C_{21}x_{1,k}+C_{22}x_{2,k}+D_{21}w_k\end{cases} \tag{6.86}$$

Then, if A_2 is nonsingular (Assumption 6.3.5) we can solve for x_2 in equation (6.86(b)) to get

$$\bar{x}_2=-A_2^{-1}(A_{21}x_1+B_{21}w)$$

and the filter (6.83) takes the following form

$$\mathbf{F}_{DrN4}^{ld}\begin{cases}\breve{x}_{1,k+1} &= \breve{A}_1\breve{x}_{1,k}+\breve{B}_{11}\breve{w}_1^\star+\breve{L}[y_k-\breve{C}_{21}\breve{x}_{1,k}-D_{21}\breve{w}_1^\star];\quad \breve{x}_1(k_0)=0\\ \breve{z}_k &= y_k-\breve{C}_{21}\breve{x}_{1,k}.\end{cases} \tag{6.87}$$

where $\breve{A}_1=(A_1-A_{12}A_2^{-1}A_{21})$, $\breve{B}_{11}=(B_{11}-A_{12}A_2^{-1}B_{21})$, $\breve{C}_{21}=(C_{21}-C_{22}A_2^{-1}A_{21})$.

Then, we have the following corollary.

Corollary 6.3.2. *Consider the linear descriptor system (6.64) and the $\mathcal{H}_\infty$-filtering problem for this system. Suppose the plant $\mathbf{P}_D^{ld}$ is locally asymptotically stable about the equilibrium-point $x = 0$, Assumption 6.3.5, Assumption 6.3.6 hold, and the plant is observable. Suppose further, for some $\gamma > 0$, there exist symmetric positive-semidefinite matrices $\breve{P} \in \Re^{q\times q}$, $\breve{Q} \in \Re^{m\times m}$, and a matrix $\breve{L} \in \Re^{n\times m}$, satisfying the LMIs:*

$$\begin{bmatrix} 3\breve{A}_1^T\breve{P}\breve{A}_1 - \breve{P} + (1-\gamma^2)\breve{C}_2^T\breve{C}_{21}^T & \breve{B}_{11}^T\breve{P}\breve{A}_1 & (1-\gamma^2)\breve{C}_{21}^T & 0 \\ \breve{A}_1^T\breve{P}\tilde{B}_{11} & -\gamma^2 I & 0 & 0 \\ (1-\gamma^2)\breve{C}_{21} & 0 & \breve{Q} - I & 0 \\ 0 & 0 & 0 & -\breve{Q} \end{bmatrix} \le 0 \tag{6.88}$$

$$\begin{bmatrix} 0 & \frac{1}{2}(\breve{A}_1^T\breve{P}\breve{L} - \gamma^2\breve{C}_{21}^T) \\ \frac{1}{2}(\breve{A}_1^T\breve{P}\breve{L} - \gamma^2\breve{C}_{21}^T)^T & (1-\delta_3)I \end{bmatrix} \le 0 \tag{6.89}$$

for some number $\delta_3 \geq 1$. Then, the filter (6.87) solves the $\mathcal{H}_\infty$-filtering problem for the system.

Proof: Take $\breve{V}(\breve{x}) = \frac{1}{2}(\breve{x}_1^T\breve{P}\breve{x}_1 + y^T\breve{Q}y)$ and apply the result of the Theorem. □

6.4 Examples

Consider the following simple nonlinear differential-algebraic system:

$$\begin{aligned} x_{1,k+1} &= x_{1,k}^{1/3} + x_{2,k}^{2/5} && (6.90) \\ 0 &= x_{1,k} + x_{2,k} && (6.91) \\ y_k &= x_{1,k} + x_{2,k} + w_k. && (6.92) \end{aligned}$$

where $w \in \ell_2$. A singular filter of the form $\mathbf{F}_{DS1}^a$ (6.49) presented in Subsection 6.3.2 can be designed. It can be checked that, the system is locally zero-input observable, and with $\gamma = 1$for $\gamma = 1$, the function $\hat{V}(\hat{x}) = \frac{1}{2}(\hat{x}_1^2 + \hat{x}_2^2 + y^2)$, solves the DHJIE (6.59) for the system.

Subsequently, we calculate the gain of the filter as

$$\hat{l}_1(\hat{x}_k, y_k) = -\frac{(y_k - \hat{x}_{1,k} - \hat{x}_{2,k})}{\hat{x}_{1,k}^{1/3} + \hat{x}_{2,k}^{2/5}},$$

where $\hat{l}_1$ is set equal zero if $|\hat{x}_{1,k}^{1/3} + \hat{x}_{2,k}^{2/5}| < \epsilon$ (ϵ small) to avoid a singularity. Thus, $x_{1,k}$ can be estimated with the filter, while $x_{2,k}$ can be estimated from $\hat{x}_{2,k} = -\hat{x}_{1,k}$.

Similarly, a normal filter of the form (6.83) can be designed. It can be checked that, Assumption 6.3.5 is satisfied, and the function $\breve{V}(\breve{x}) = \frac{1}{2}(\breve{x}_1^2 + y^2)$ solves the DHJIE (6.80) for the system. Consequently, we can also calculate the filter gain as

$$\breve{l}_1(\breve{x}_k, y_k) = -\frac{(y_k - \hat{x}_{1,k} - \breve{x}_{2,k})}{\breve{x}_{1,k}^{1/3} + \breve{x}_{1,k}^{2/5}}$$

and again $\breve{l}_1$ is set equal zero if $|\breve{x}_{1,k}^{1/3} + \breve{x}_{1,k}^{2/5}| < \epsilon$ (ϵ small) to avoid a singularity.

6.5 Conclusion

In this chapter, we have presented a solution to the $\mathcal{H}_\infty$ filtering problem for affine nonlinear descriptor systems. Both in continuous-time, and in discrete-time. Two types of filters have been proposed; namely, singular and normal filters. Reduced-order normal filters have also been presented for the case of standard systems. Sufficient conditions for the solvability of the problem using each type of filter are given in terms of HJIEs and DHJIEs. The results have also been specialized to linear systems, in which case the conditions reduce to a system of LMIs which are computationally efficient to solve. The problem for a nonconstant singular recursion matrix has also been discussed, and finally, examples have been presented to demonstrate the approach.

clearpage

SUMMARY AND CONCLUSION

In this Dissertation, we have reviewed the historical development of estimation theory from Gauss's least squares method to the Kalman-Bucy theory and finally to the Stratonovich-Kushner theory. We have also summarized most of the major approaches that have been developed for linear dynamic systems, including the minimum-variance method, the maximum likelihood method and the Bayesian approaches. Finally, we have also discussed the extensions of the above approaches to nonlinear dynamic systems including the extended Kalman filter, the Stratonovich and Kushner filters as well as the maximum likelihood recursive nonlinear filters and Baysian nonlinear filters.

On the other hand, the contribution of the Dissertation is mainly to develop $\mathcal{H}_2$ and $\mathcal{H}_\infty$ approaches to filtering for nonlinear singular systems. These approaches which are deterministic are much easier to derive and the filters developed are simpler to inplement. The filters derived are also finite-dimensional, as opposed to the statistical methods which lead to infinite-dimensional filters and evolution equations such as the Stratonovich equation, the Kushner equation, and the Wong and Zakai equation, which have no exact solutions and neither computationally tractable numerical solutions. They rely on finding smooth solutions of certain Hamilton-Jacobi equations which can be found using polynomial approximations or other methods (Aliyu, 2003), (Al-Tamimi, 2008), (Abukhalaf, 2006), (Huang, 1999).

Further, to summarize the results, for singularly-perturbed systems, we have presented $\mathcal{H}_2$ and $\mathcal{H}_\infty$-decomposition, aggregate and reduced-order filters, both in continuous-time and discrete-time. While for descriptor systems, we have presented similarly $\mathcal{H}_2$ and $\mathcal{H}_\infty$-singular and normal filters in both continuous-time and discrete-time. Reduced-order filters have also been considered. Some simulation results have also been presented to validate the approaches.

The $\mathcal{H}_2$ filters are useful when the system and measurement noise is fairly known and can be modeled as stationary Gaussian white-noise processes with certain covariances. Whereas the $\mathcal{H}_\infty$ filters are more useful when the system and measurement noise are generally unknown, but can be assumed to be $\mathcal{L}_2$-bounded. They also have better robustness against other forms of disturbances than the $\mathcal{H}_2$ filters. In both cases we have pursued deterministic approaches to the filter design as apposed to stochastic approaches.

By-and-large the Dissertation represents the first successful attempt to use Hamilton-Jacobi theory to solve the filtering problem for affine nonlinear systems. Earlier not so successful attempts (Mortenson, 1968), (Berman, 1996), (Nguang, 1996), (Shaked, 1995) have led to very complicated Hamilton-Jacobi equations involving a rank-3 tensor (Mortenson, 1968) and gain matrices that require the original state information (Berman, 1996), (Nguang, 1996), (Shaked, 1995) which are practically unrealizable. On the other hand, we have avoided both of these two problems by not using an error vector $e = x - \hat{x}$ in our design, and also by using determinstic techniques. We have attempted to address all the research objectives that we set out in Chapter 2, but we believe that improvements can still be made on the results that we have achieved, especially in finding efficient ways to solve the Hamilton-Jacobi equations. Future efforts will also consider ℓ_1 filter design approaches which have the tendency to suppress persistent bounded disturbances as well.

REFERENCES

Aganovic, Z., Gajic, Z., and Shen, X., (1996), "New Method for Optimal Control and Filtering of Weakly Coupled Linear Discrete Stochastic Systems," Automatica, Vol. 32, No. 1, pp. 83-88.

Aliyu, M. D. S., (2003) "An Approach for Solving the Hamilton-Jacobi-Isaac Equations Arising in Nonlinear $\mathcal{H}_\infty$-Control," Automatica, Vol. 39, No. 5, pp. 877-884.

Aliyu, M. D. S., and Boukas, E. K., (2009) "Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ Nonlinear Filtering," Int. J. of Robust and Nonlinear Control, Vol. 19, No. 4, pp. 394-417.

Aliyu, M. D. S., and Boukas, E. K., (2011), "$\mathcal{H}_2$ Filtering for Discrete-time Singularly-Perturbed Nonlinear systems," To appear in IEEE Trans, Circuits and Systems-I: Fundamental Theory and Appl.

Aliyu, M. D. S., and Boukas, E. K., (2011), "$\mathcal{H}_\infty$ Filtering for Singularly-Perturbed Nonlinear systems," Int. J. of Robust and Nonlinear Control, Vo. 21, pp. 218-236.

Aliyu, M. D. S. and Boukas, E. K., (2011), "$\mathcal{H}_2$ Filtering for Singularly-Perturbed Nonlinear Systems," Accepted in IET-Control Thry and Appl.

Al-Tamimi, A., Lewis, F. L., and Khalaf, M. A., (2008), "Discrete-time Nonlinear HJB Solution Using Approximate Dynamic Programming: Convergence Proof," IEEE Trans. Sys, Man and Cybern, Vol. 38, No. 4, pp. 943-949.

Abu-Khalaf, M., Lewis, F. L., Huang J., (2006), "Policy Iterations on the Hamilton-Jacobi-Isaacs Equation for $\mathcal{H}_\infty$ State Feedback Control With Input Saturation," IEEE Trans. Autom Control, Vol. 51, No. 12, pp. 1989-1995.

Anderson, B. D. O., and Moore, J. B., (1979). *Optimal Filtering.* Prentice Hall, NJ.

Arasaratnam, I., Haykin, S. and Elliot, R. J., (2007), "Discrete-time Nonlinear Filtering Algorithms Using Gauss-Hermite Quadrature," Proc. of IEEE, Vol. 95, No. 5., pp. 953-977.

Arasaratnam, I. and Haykin, S., (2009),"Cubature Kalman Filters," IEEE Trans. Autom Control, Vol. 95, No. 54, No. 6, pp. 1254-1269.

Assawinchaichote, W., Nguang, S. K., and Shi, P., (2004),"$\mathcal{H}_\infty$ Output-feedback Control Design for Nonlinear Singularly-perturbed Systems: An LMI Approach," Automatica. Vol. 40, No. 12. 2147-2152.

Assawinchaichote, W., and Nguang, S. K., (2004), "$\mathcal{H}_\infty$ Filtering for Fuzzy Nonlinear Singularly-perturbed Systems with Pole Placement Constraints: An LMI Approach," IEEE Trans on Signal Processing, Vol. 52, pp. 1659-1667.

Assawinchaichote, W., and Nguang, S. K., (2007), "Robust $\mathcal{H}_\infty$ Fuzzy Filter Design for Uncertain Nonlinear Singularly-perturbed Systems with Markovian Jumps: An LMI Approach," J. of Information Sciences, Vol. 177, pp. 1699-1714.

Bagchi, A., (1993). *Optimal Control of Stochastic Systems*. Prentice Hall, UK.

Basar, T. and Bernhard, P. *H_∞-Optimal Control and Related Minimax Design Problems*, (1995). Systems and Control Foundations and Applications, Second Edition, Birkhauser, Boston.

Basar, T., and Olsder, G. J., (1982). *Dynamic Noncooperative Game Theory*. Mathematics in Science and Engineering, Vol. 160, First Edition, Academic Press, UK.

Berman, N. and Shaked, U., (1975), "$\mathcal{H}_\infty$ Nonlinear Filtering," Int. J. of Robust and Nonlinear Control, Vol. 6, pp. 281-296.

Boothby, W. M., (1975). *An Introduction to Differentiable manifolds and Riemannian Geometry*. Academic Press, UK edition.

Boulkroune B., Darouach, M., and Zasadzinski, M., (2010), "Optimal Unbiased Reduced-order Filtering for Discrete-time Descriptor Systems via LMI," IET Control Thry and Appl., Vol. 4, No. 3, pp. 339-350.

Bouyekhf, R., and Moudhni, A. E., (1996), "Reduced-Order Modeling of Two-time Scale Discrete-time Nonlinear systems," J. of Franklin Inst., Vol. 334B. No. 4, pp. 499-512.

Bouyekhf, R., and Moudhni, A. E., (1997),"On the Analysis of Discrete-time Singularly Perturbed Nonlinear systems: Application to the study of Stability Properties," J. of Franklin Inst., Vol. 334B. No. 2, pp. 199-212.

Bouyekhf R. and Moudhni A. E., (1998),"Stabilization and Regulation of a Class of Nonlinear Singularly Perturbed Discrete-time Systems," J. of Franklin Inst., Vol. 334B. No. 5, pp. 963-982.

Boyd S., El-Ghaoui, L., Feron, E., and Balakrishnan, V., (1994). *Linear Matrix Inequalities in Systems and Control Theory.* SIAM.

Bucy, R. S., (1970), "Linear and Nonlinear Filtering," IEEE Proc., Vol. 58, No. 6, pp. 854-864.

Bucy, R. S., (1967), "Nonlinear Filtering Theory," IEEE Trans Autom Control, Vol. 10, pp. 198.

Bucy, R. S., (1967), "Optimal Filtering for Corelated Noise," J. of Mathematical Analysis, Vol. 20, pp. 1-8.

Chang, K. W., (1972), "Singular Pertubation of a General Boundary-value Problem," SIAM J. Math Anal. Vol. 3, pp. 520-526.

Chow, J. H., and Kokotovic, P. V., (1976),"A Decomposition of Near-Optimal Regulators for Systems with Slow and Fast Modes," IEEE Trans Autom Control, Vol. 21, pp. 701-705.

Crassidis, J. L., and Junkins, J. L., (2004). *Optimal Estimation of Dynamic Systems.* Chapman and Hall, Applied Maths and Nonlinear Science Series.

Dai, L., (1989),"Filtering and LQG Problem for Discrete-time Stochastic Singular Systems," IEEE Trans Autom Control, Vol. 34, No. 10, pp. 1105-1108.

Dai., L., (1989). *Singular Control Systems.* Springer Verlag, Lecture Notes in Control and Information Sciences, Vol. 118, Germany.

Dai, L., (1988), "Observers for Discrete Singular Systems," IEEE Trans Autom Control, Vol. 33, No. 2, pp. 187-191.

Darouach M., and Boutayeb, M. (1995), "Design of Observers for Descriptor Systems," IEEE Trans Autom Control, Vol. 40, No. 7, pp. 1323-1327.

Darouach, M., Zasadzinski, M., and Hayar, M., (1996), "Reduced-Order Observer Design for Descriptor Systems with Unknown Inputs," IEEE Trans Autom Control, Vol. 41, No. 7, pp. 1068-1072.

Darouach, M., and Boutat-Baddas, L., (2008), "Observers for a Class of Nonlinear Singular Systems," IEEE Trans Autom Control, Vol. 53, No. 11, pp. 2627-2633.

Darouach, M., and Boutayyeb, M., and Zasadzinski M., (1997), "Kalman Filtering for Continuous Descriptor Singular," Proc. American Control Conf., Albuqueque, New Mexico pp. 2108-2112.

Darouach, M. and Zasadzinski, M. (2009), "Moving-Horizon State-Estimation for Linear Discrete-time Singular Systems," Systems and Control Lett., Vol. 58, pp. 436-444.

Daum F., (2005), "Nonlinear Filters: Beyond the Kalman Filter," IEEE A& E Magazine, Vol. 20, No. 8, August, pp. 57-69.

El-Tohami M., Lovass-Nagy, V., and Mukundan, R., (1984), "Design of Observers for Time-varying Discrete-time Descriptor Systems," Int. J. of Control, Vol. 46, pp. 841-848.

El-Tohami, M., Lovass-Nagy, V., and Powers, D. L., (1984), "On Input Observers for Generalized State-space Systems," Int. J. of Control, Vol. 40, No. 5, pp. 903-922.

El-Tohami, M., Lovass-Nagy V., and Mukundan, R., (1983), "On the Design of Observers for Generalized State-space Systems Using Singular-Value Decomposition," Int. J. of Control, Vol. 38, No. 3, pp. 673-683.

Fahmy, M. M., and O'Reilly, J., (1989),"Observers for Descriptor Systems," Int. J. of Control, Vol. 49, No. 6, pp. 2013-2028.

Feng, Y., Anderson, B. D. O., and Rotkowitz, M., (2009),"A Game Theoretic Algorithm to Compute Local Stabilizing Solutions of HJBI Equations in Nonlinear $\mathcal{H}_\infty$ Control," Automatica, Vol. 45, pp. 881-888.

Gajic Z., and Lim, M.-T., (1994),“A New Filtering Method for Linear Singularly Perturbed Systems,” IEEE Trans Autom Control, Vol. 39, No. 9, pp. 1952-1955.

Gao, Z. and Ho, D. W. C., (2004), “Proportional Multiple-Integral Observer Design for Descriptor Systems with Measurement Output Disturbances,”, IEE Proc Control thry Appl., Vol. 151, No. 3, pp. 279-288.

Gordon, N. J., Salmond, D. J. and Smith, A. F. M., (1993), “Novel Approach to Nonlinear/Non-Gaussian Baysian State Estimation,” IEE Proc. Part F., Vol. 40, No. 2, pp. 107-113.

Grewal, M. S., and Andrews A. P. (1193). *Kalman Filtering: Theory and Practice*. Prentice Hall Information and Systems Science Series, NJ.

Guillard H., Monaco S., and Normand-Cyrot D., (1996), “On the $\mathcal{H}_\infty$ Control of Discrete-time Nonlinear Systems,” Int. J. of Robust and Nonlinear Control, Vol. 6, pp. 633-643.

Haddad, A. H., (1976), “Linear Filtering of Singularly Perturbed Systems,” IEEE Trans on Autom Control, Vol. 21 pp. 515-519.

Haddad, A. H., and Kokotovic, P., (1977), “Stochastic Control of Linear Singularly Perturbed Systems,” IEEE Trans on Autom Control, Vol. 2, pp. 815-821.

Henrikson, R., (1982), “The Truncated Second-Order Nonlinear Filter Revisited,” IEEE Trans. Autom Control, Vol. 27, N0. 1, Feb. pp. 247-251.

Ho, Y. C. and Lee, R. C. K., (1964), “A Bayesian Approach to Problems in Stochastic Estimation and Control,” IEEE Trans. Autom Control, Vol. 9, pp. 333-339.

Hong G., and Dong, J., (2008), “$\mathcal{H}_\infty$ Filtering for Fuzzy Singularly-Perturbed Systems,” IEEE Trans on Systems, Man and Cybernetics-Part B, Vol. 38, No 5, 1371-1389.

Hou M., and Muller, P. C., (1995),“Design of a Class of Luenberger Observers for Descriptor Systems,” IEEE Trans Autom Control, Vol. 40, No. 1, pp. 133-136, January.

Huang, J., (1999), “An Algorithm to Solve the Discrete-time HJI Equation Arising in the $\mathcal{L}_2$-gain Optimization Problem,” Int. J. Control, Vol. 72, No. 1, pp. 49-57.

Ishihara, J. Y., Terra, M. H., and Campos, J. C. T., (2006),“Robust Kalman Filter for Descriptor Systems,” IEEE Trans Autom Control, Vol. 51, No. 8, pp. 1354-1358.

Ishihara, J. Y., Terra, M. H., and Espinoza, B. M., (2009), “$\mathcal{H}_\infty$ Filtering for Rectangular Descriptor Systems,” Automatica, Vol. 45, pp. 1743-1748.

Jazwinski, A. H., (1998). *Stochastic Processes and Filtering Theory*. Dover Publ., NY, USA.

Julier, S., Uhlmann, J., and Durrant-Whyte, H. F., (2000),“A New Method for the Nonlinear Transformation of the Means and Covariances in Filters and Estimators,” IEEE Trans. Autom Control, Vol. 45, No. 3, pp. 477-482.

Kalman, R. E., (1960),“A New Approach to Linear Filtering and Prediction Problems,” Trans. ASME J. Basic Engineering, Series D, Vol. 82, 35-45.

Kalman, R. E., and Bucy, R. S., (1961), “New Results in Linear Filtering and Prediction Theory,” Trans. ASME J. Basic Engineering, Series D, Vol. 83, 95-108.

Khalil, H. K., (1985). *Nonlinear Systems*. Mcmillan, USA.

Khalil, H. K., and Gajic, Z., (1984), “Near-optimum Regulators for Stochastic Linear Singularly Perturbed Systems,” IEEE Trans. Automat Control, Vol. 29, pp. 531-541.

B-S. Kim, Y-J. Kim Y-J., and M-T. Lim, (2002), “LQG Control for Nonstandard Singularly Perturbed Discrete-time Systems,” Proc. IEEE Conf. on Dec. and Control, Las Vegas, Nevada, pp. 3736-3741.

Koenig, M. D. and Mammar, S., (1995),“Design of Proportional-Integral Observer for Unknown Input Descriptor Systems,” IEEE Trans Autom Control, Vol. 47, No. 7, pp. 2057-2062.

Kokotovic, P., Khalil, H. K., and O’Reilly J. (1986). *Singular Pertubation Methods in Control: Analysis and Design*. Academic Press, USA.

Kolmogorov, A. N., (1949), “Interpolation and Extrapolation,” Bull. de l’Académie des Sciences de USSR, Ser. Math, 5, pp. 3-41, 1949.

Kushner, H. J., (1967), “Nonlinear Filtering: The Exact Dynamical Equation Satisfied by the Conditional Mode,” IEEE Trans. Autom Control, Vol. 12, No. 3, pp. 262-267.

Kushner, H. J., (1964), “On the Differential Equation Satisfied by the Conditional Probability Densities of Markov Processes with Applications,” SIAM J. on Control, Ser A, Vol. 2, No. 1, pp. 106-119.

Kushner, H. J., (1964),“On the Dynamical Equations of Conditional Probability Density Functions, with Applications to Optimal Stochastic Control Theory,” J. of Mathematical Analysis and Applications, Vol. 8, pp. 332-344.

Kushner, H. J., (1967),“Approximations to Optimal Nonlinear Filters,” IEEE Trans. Autom Control, Vol. 12, No. 5, pp. 546-556.

Kushner, H. J., (1974),“An Approach to Useful but Nonoptimal Nonlinear Filters,” IEEE Trans. Autom Control, Vol. 19, pp. 398-401.

Kushner, H. J., (2000), “A Nonlinear Filtering Algorithm Based on Approximation of the Conditional Distribution,” IEEE Trans. Autom Control, Vol. 45, No. 3, pp. 580-585.

Lee, J. S. and Tung, F. F., (1970),“Nonlinear Filtering Techniques with Application to Strapdown Computation,” IEEE Trans. Autom Control, Vol. 15, No. 1, pp. 74-81.

Lim, M. T., Gajic, Z., and Shen, X., (1996), “New Methods for Optimal Control and Filtering of Weakly Coupled Linear Discrete Stochastic Systems,” Proc, American Control Conf., Seattle Wash., June pp. 534-538.

Lim, M. T. and Gajic, Z., (2000) “Reduced-order $\mathcal{H}_\infty$ Optimal Filtering for Systems with Slow and Fast Modes,” IEEE Trans on Circuits and Systems, Vol. 47, No. 2, pp. 250-254.

Marcus, S. I., (1979),“Optimal Nonlinear Estimation for a Class of Discrete-time stochastic Systems,” IEEE Trans Autom Control, Vol. 24, No. 2, pp. 297-302.

N. Minamide, (1989), “Design of Observers for descriptor systems using a Descriptor standard Form,” Int. J. control, Vol. 50, No. 6, pp. 2141-2149.

R. E. Mortensen, (1968), “Maximum-Likelihood Recursive Nonlinear Filtering,” J. of Optimiz. Thry and Appl. (JOTA), Vol. 2, No. 6, pp. 386-394.

H. Mukaidani, (2003), “Near-Optimal Filters for Multiparameter Singularly Perturbed Linear Systems,” IEEE Trans Circuits and Systems-I: Fundamental Thry and Appl., Vol. 50, No. 5, pp. 717-721.

Newcomb, R. W., (1981), “The semistate Description of Nonlinear Time-variable Circuits,” IEEE Trans Cicuits and Systems, Vol. 28, No. 1, pp. 62-71.

Newcomb, R. W., and Goknar I. C., (1981), “On the State Description of Nonlinear Networks-Further Results,” Proc. IEEE, Vol. 69, No. 5, pp. 659-660.

Netto, M. L. A., Gimeno, L., and Mendes, M. J., (1978), “On the optimal and Sub-optimal Nonlinear Filtering Problem for Discrete-time Systems,” IEEE Trans. Autom. Control, Vol. 23, No. 6, pp. 1062-1067.

Nikoukhah, R., Campbell, S. L., and Delebecque, F., (1999), “Kalman Filtering for General Discrete-time Linear Systems,” IEEE Trans Autom Control, Vol. 44, No. 10, pp. 1829-1839.

Nikoukhah, R., Willsky, A. L., and Levy, B. C., (1992), “Kalman Filtering and Riccati Equations for Descriptor Systems,” IEEE Trans Autom Control, Vol. 37, No. 9, September pp. 1325-1342.

Nguang, S. K., and Fu, M., (1996), “Robust Nonlinear Filtering,” Automatica, Vol. 30, No. 7, pp. 1195-1199.

Ohmatsu, S., Tomita, Y., and Soeda, T., (1976),“Optimal Filtering for Discrete-time Nonlinear Systems,” IEEE Trans. Autom Control, Vol. 21, pp. 116-118.

Ozcaldiran, K., Fountain, D. W., and Lewis, F. L., (1992),“Some Generalized Notions of Observability,” IEEE Trans Autom Control, Vol. 37, No. 6, pp. 856-860.

Paraskevopoulos, P. N., and Koumboulis, F. N., (1992), “Observers for Singular Systems,” IEEE Trans Autom Control, Vol. 37, No. 8, pp. 1211-1215.

Prljaca, N., and Gajic, Z., (2008), "A Method for Optimal Control and Filtering of Multitime-scale Singularly Perturbed Stochastic Systems," Automatica, Vol. 44, pp. 2149-2156.

Rauch, H. E., Tung, F., and Striebel, C. T., (1965), "Maximum-Likelihood Estimates of Linear Dynamic Systems," AIAA J., Vol. 3, pp. 1445-1450.

Sadjadi-Azmi, M. R., and Khorasani, K., (1990), "Reduced-Order Strip Kalman Filtering Using Singular Pertubation Method," IEEE Transactions Circuits and Systems, Vol. 37, No. 2, pp. 284-284.

Sakamoto, N., and Van der Schaft, A. J., (2008),"Analytical Approximation Methods for the Stabilizing Solution of the Hamilton-Jacobi Equation," IEEE Trans. Autom Control, Vol. 53, No 10, pp. 2335-2350.

Sastry, S., (1999).*Nonlinear Systems: Analysis, Stability and Control*. Springer Verlag NY.

Sebald, A. V., and Haddad, A. H., (1978), "State Estimation for Singularly-Perturbed Systems with Uncertain Pertubation Parameter," IEEE Trans Autom Control, Vol. 23, No. 3, pp. 464-469.

Sen, S., and Datta, K. B., (1993), "Stability Bounds for Singularly Perturbed Systems," IEEE Trans Autom Control, Vol. 38, No. 2, pp. 302-304.

Shaked, U., and Berman, N., (1995), "$\mathcal{H}_\infty$ Nonlinear Filtering of Discrete-Time Processes," IEEE Trans. Signal Process, Vol. 43, No. 9, pp. 2205-2209.

Shen, X., Rao, M., and Ying, Y., (1993), "Decomposition Method for Solving Kalman filter gains in Singularly Perturbed Systems," Optimal Control Appl. and Methods, Vol 14, No. 1, pp. 1-73.

Shen, X., and L. Deng, L., (1996), "Decomposition Solution of $\mathcal{H}_\infty$ Filter Gain in Singularly Perturbed Systems," Signal Processing, Vol. 55, pp. 313-320.

Sorensen, H. W., (1985). *Kalman Filtering: Theory and Application*. Editor, IEEE Press, NY.

Stratonovich, R. L., (1960), “Application of the Theory of Markoff Processes in Optimal Signal detection,” Radio engineering Electron Phy., Vol. 1, pp. 1-19.

Stratonovich, R. L., (1966), “A New form of Representing Stochastic Integrals and Equations,” SIAM J. on Control, Vol. 4, pp. 362-371.

Sun, S-L., and Ma, J., (2007), “Optimal Filtering and Smoothing for Discrete-time Stochastic Singular systems,” Signal Processing, Vol. 87, pp. 189-201.

Uetake, Y., (1989), “Pole Assignment and Observer Design for Continuous Descriptor Systems,” Int. J. Control, Vol. 50, No. 1, pp. 89-96.

Vidyasagar, M., (1993). *Nonlinear Systems Analysis*. Prentice Hall, NJ, 1993.

Wang, H.-S., and Yung, C.-F., (2008),“Discrete-time $\mathcal{H}_\infty$ Control Problem for Nonlinear Descriptor Systems,” IEEE Trans Autom Control, Vol. 53, No. 4, pp. 1051-1057.

Wiener, N., (1949). *Extrapolation, Interpolation, & Smoothing of Stationary Time Series*. The MIT Press, Cambridge, Mass, 1949.

Wong, E., and Zakai, M., (1965),“On the Convergence of Ordinary Integrals to Stochastic Integrals,” Annals of Mathematical Statistics, 36, pp. 1560-1564.

Wonham, W. M., (1963), “Stochastic Problems in Optimal Control,” IEEE Intern. Conv. Record, Vol. 11, pp. 114-124.

Wonham, W. M., (1965), “Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering,” SIAM J. on Control, Ser A, Vol. 2, No. 3, pp. 347-369.

Xie, L., De Souza, C. E., and Wang, Y., (1996), “Robust Filtering for a Class of Discrete-Time Nonlinear Systems: An $\mathcal{H}_\infty$ Approach,” Int. J. of Robust and Nonlinear Control, Vol. 6, pp. 287-312.

Xu, S., Lam, J., Zhou, Y., (2007), “$\mathcal{H}_\infty$ Filtering for Singular Systems,” Systems & Control, Lett. Vol. 56, pp. 48-57.

Xu, S., Lam, J., (2003), "Reduced-Order $\mathcal{H}_\infty$ Filtering for Singular Systems," IEEE Trans. Autom Control, Vo. 48, pp. 2217-222.

Xu, S., Lam, J. and Zhou, Y., (2003), "$\mathcal{H}_\infty$ Filtering for Singular Systems," IEEE Trans Autom Control, Vol. 48, No. 12, pp. 2217-2222.

Yang, G. H., and Dong, J., (2008), "$\mathcal{H}_\infty$ Filtering for Fuzzy Singularly-Perturbed Systems," IEEE Trans on Systems, Man and Cybernetics-Part B, Vol. 38, No 5, 1371-1389.

Young, K.-K. D., Kokotovic, P. V., and Utkin, V. I., (1977), "Singular-Perturbation Analysis of High-gain Feedback Systems," IEEE Trans Autom Control, Vol. 22, No. 6, pp. 931-938.

Zakai, M., (1969), "On the Optimal Filtering of Diffusion Processes," Wahrschein Likehkeitstheorie verw gebiet, Vol. 11, pp. 230-243.

Zhang, W. H., Chen, B. S., and Tseng, C. C., (2005), "Robust $\mathcal{H}_\infty$ Filtering for Nonlinear Stochastic Systems," IEEE Trans on Signal Processing, Vol. 53, Part 1, No. 2, pp. 589-598.

Zhou, Y., and Li, J., (2008), "Reduced-Order $\mathcal{L}_2 - \mathcal{L}_\infty$ Filtering for Singular Systems: An LMI Approach," IET Proc Control Thry Appl., Vol. 2, No. 3, pp. 228-238.

Zhou, Y., and Li, J., (2008),"Energy to Peak Filtering for Singular systems: The Discrete-time Case," IET Proc Control Thry Appl., Vol. 2, No. 9, pp. 773-781.

Zimmer, G. and Meier, J., (1997), "On Observing Nonlinear Descriptor Systems," Sys. and Control Lett. , Vol. 32, pp. 43-48.